MONASH University

Neural Correlates of Speed Processing in the Primate Visual Area MT

Amanda Davies

Bachelor of Behavioural Neuroscience (Hons)

A thesis submitted for the degree of *Doctor of Philosophy* at

Monash University in 2017

Department of Physiology, Faculty of Medicine, Nursing and Health Sciences

Copyright notice

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

Abstract

One of the crucial functions of the visual system is the ability to estimate the speed of a moving object irrespective of its size or texture. Previous studies have suggested that the neuronal coding of speed is ambiguous, with most neurons changing their speed tuning depending on the spatial frequency of a visual pattern. This form of response is considered temporal frequency-tuned, where the neurons respond to the number of black/white cycles per second. For a reliable interpretation of speed, neurons are instead required to respond to a consistent speed regardless of the spatial frequency of the stimulus; namely, they respond to angular velocity (degrees per second). It is hypothesised that this could be an effect of the drifting (grating) stimulus that has previously been employed, which does not reflect characteristics of natural visual scenes. This thesis investigates how the introduction of a stimulus trajectory affects the response pattern of a large majority of neurons in the middle temporal area (MT), a cortical area specialised for motion processing.

Two areas on the motion processing pathway, MT and the primary visual area (V1), were investigated using marmoset monkeys. As a crucial part of the motion processing pathway, MT is a likely candidate for a population of speed-tuned neurons, while V1 was investigated due to the large input it provides to MT; allowing us to determine whether MT speed-tuning is derived from V1 input. Electrophysiological techniques were used to investigate neuronal responses in V1 and MT to a range of stimuli. A model was also designed according to the commonly used model of linear summation, allowing us to determine if our results could be predicted by our model.

For three different stimulus motion conditions, dependency of speed tuning on the spatial frequency of sine waves was established for both V1 and MT. The three stimulus conditions consisted of: the Standard mode (a fixed envelope, drifting Gaussian sine wave which replicates that of the stimulus used in previous studies); the Moving mode (a fixed Gaussian

sine wave within a moving envelope, which provides a predictable trajectory that crosses the cell's receptive field); and the Waxing/Waning mode (a fixed envelope, drifting Gaussian sine wave with a varying luminance pattern which acted as a bridge between the previous two modes).

Within MT, the use of the Moving mode stimulus revealed a much greater proportion of speed-tuned neurons (n = 50, 45%) when compared to the two other stimuli where the sine wave envelope remained stationary over the receptive field. When tested in the primary visual area, even using Moving mode stimuli, there was only a small population of speed tuned neurons (n = 6, 7%), suggesting that the spatial frequency independent speed-tuning found in MT is not likely to be derived solely from V1 inputs. Modelling of the data showed that the results we found could not be explained solely according to the current linear summation model.

Together, these results show that the type of stimulus used for probing neuronal responses to motion can have a critical effect on the detection of speed tuned neurons The results suggest that the prevalence of stimulus-invariant speed-tuning in MT has been previously underestimated, in part due to the use of stimuli that do not incorporate spatial and temporal context such as those present in natural scenes. This has shown that the brain processes speed using a more complicated mechanism than previously thought, and that future models should endeavour to account for the additional non-linear mechanism found in this thesis.

Declaration

In accordance with Monash University Doctorate Regulation 17/ Doctor of Philosophy and Master of Philosophy regulations the following declarations are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and beliefs, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

All experiments described in the thesis involving animals were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes, and approved by the Monash University Animal Ethics Experimentation Committee.

The core theme of this thesis is to examine the effects of a predictable trajectory on the processing of speed within the visual system. The idea, development and writing up of all the chapters in this thesis were the principal responsibility of myself, the candidate, working within the Department of Physiology, Monash University, under the supervision of Professor Marcello GP Rosa and Dr Hsin-Hao Yu.

Amanda Davies

Date

Abstracts and Publications

International Meetings

2015 45th SFN Annual Neuroscience Meeting, Chicago, USA

AJ Davies, MGP Rosa, HH Yu, (2015) '*Context-dependent robust coding of stimulus speed in primate extrastriate cortex.'*

2015 11th Asia-Pacific Conference on Vision, Singapore, Singapore

AJ Davies, T Chaplin, MGP Rosa, HH Yu, (2015) '*Context-dependent robust coding of stimulus speed in primate extrastriate cortex.'*

2013 33rd Annual Meeting of the Australian Neuroscience Society, Melbourne, Australia

AJ Davies, HH YU, TA Chaplin, MGP Rosa, (2013) '*Contextual effects of speed-tuning of neurons in the middle temporal area (MT)'*

Workshops

2013 IBRO/APRC Advanced Neuroscience Imaging Course-Program, Clayton, Victoria

Publications

Yu HH, Chaplin TA, Davies AJ, Verma R, Rosa MGP (2012) A specialized area in limbic cortex for fast analysis of peripheral vision. *Current Biology, 22,* 1351-1357.

Davies AJ, Chaplin TA, Rosa MGP, Yu HH (2016) Natural motion trajectory enhances the coding of speed in primate extrastriate cortex. *Scientific Reports*, 6, 19739.

Acknowledgements

This thesis was such a large undertaking that it would have been impossible to get it done without the support of my friends and family. They were there to constantly put up with my stress as I worried about getting it done fast enough, or to help by reading through things even if they didn't have any idea what I was writing about. Knowing that that support is there behind me, that I have their best wishes, often helps you to keep pushing through even when you begin to doubt yourself.

I also wanted to thank Sharon Flecknoe and Kate Carroll, who were absolutely amazing to give an untested PhD student an opportunity to teach in their fantastic new subject. Even more, the constant support and understanding as I juggled my new job and my commitments to my PhD and thesis are one of the things that made this happen.

But at the end of the day my greatest thanks need to go to the people that helped me begin my thesis. To Marcello for taking a risk on a student that may not have had the highest undergraduate marks, but had a passion for the project he was offering. The sheer depth and breadth of his knowledge continues to amaze me, as well as his passion for what he does. To Ramesh for convincing him I was worth the chance, but also for continuing to support me in my passion for teaching, I wouldn't be where I am if it wasn't for his help and guidance. Last but definitely not least, to Hsin-Hao, who has been my rock through this entire PhD and thesis, who always found time to help me, no matter how stressed he was. If his support of me through this increasingly stressful time is any indication, his beautiful new daughter is extremely lucky to have a father that will always be there to support her whenever she needs.

Table of Contents

1 General Introduction

Often humans take for granted their senses, considering how little conscious effort is required to perceive what is in the world around them. Yet these processes occur through extremely complex interactions between neurons located across many different areas of the brain. Arguably, one of the most important senses of primates is vision, as evidenced by the elaborate neural circuits in the primate brain devoted to the processing of visual information. Based on several decades of research, we now know that our seemingly unified perception of vision is multifaceted – whereby visual attributes such as shapes, colour, and motion are distinct phenomena, and are to a significant extent, mediated by separate neural mechanisms in the brain.

Of these phenomena, the ability to cope with visual information that changes with time is critical for survival, because we live in an intrinsically dynamic environment. In particular, the brain needs to be able to determine in which direction an object is moving, as well as how quickly, in order to be able to react appropriately. Within the middle temporal area (MT) in the primate visual cortex, a large population of motion sensitive neurons have been found, making it an area of great interest for studies of cortical motion processing. However, while the coding of motion direction has been extensively studied in area MT, the coding of motion speed has not. This will be the main question that I will address in this thesis.

This first chapter of my thesis is organised as follows. I will first describe the general organisation of the primate visual cortex (Section 1.1), and then go into more depth about the properties of area MT, where I used electrophysiological methods to study how the speeds of moving patterns are encoded by neurons in this area (Section 1.2). Area MT is part of a pathway where motion-related information is transformed at different stages. The primary visual cortex (V1) is of particular interest because it provides the majority of inputs to MT. In Section 1.3, I will describe the relationship between the two areas, and provide background information about my study of motion processing in area V1. In Section 1.4, I will explain the computational problem of encoding motion speed in the visual cortex. This background knowledge will motivate the design of a novel paradigm for studying motion processing, which will be explained in Section 1.5. My experimental results suggest a form of spatiotemporal integration of motion signals that have not been demonstrated before. Therefore, in Section 1.6 I will review classical computational theories of spatiotemporal integration of motion signals, in order to establish the theoretical framework needed to interpret my modelling results. I will then provide a brief review of my model animal, the marmoset monkey, in Section 1.7 and finish by stating the overall aims of my thesis in Section 1.8.

1.1 General Organisation of the Visual Cortex

Information within the primate visual system is processed in what is currently conceived of as a hierarchical network consisting of more than 30 areas (Felleman & Van Essen, 1991). Neurons in areas corresponding to "low" hierarchical stages process information from small regions of the visual field (their *receptive fields*), and respond selectively to simple image features (such as the presence of a vertical line). As information is then passed to successively "higher" stages in the hierarchy, we find that neurons tend to have progressively larger receptive fields, and that their responses tend to become increasingly more selective for complex combinations of features.

The first point of entry into this cortical hierarchy is the primary visual area (V1). Neurons in this area have been shown to have very small receptive fields (<1 degree in the central visual field), and are selective to simple visual features such as the local orientation of edges (Hubel & Weisel, 1959; Hubel & Weisel, 1968). Beyond V1, the visual cortex is classically divided into two 'streams' of processing (Ungerleider & Mishkin, 1982; Miskin et al., 1983; Ungerleider & Haxby, 1994; Kravitz et al., 2011). One pathway courses ventrally, interconnecting the striate and inferior temporal areas, and is associated with visual processing involving shapes, colours, textures, and object identification. The other pathway, which courses dorsally to the parietal cortex, mediates functions related to spatial relationships and motion processing. The distinction between dorsal and ventral pathways was originally suggested by lesion studies (Ungerleider & Mishkin, 1982; Miskin, Ungerleider & Macko, 1983; Ungerleider & Haxby, 1994), where macaque monkeys with parieto-occipital or infero-temporal lesions were tested in their performance on pattern and landmark discrimination tasks. Animals with lesions in the parieto-occipital region performed poorly at landmark discrimination tasks, while those with lesions of the inferior temporal region showed a decrease in performance during the pattern discrimination tasks. These

experiments suggested a distinction between the two streams: the inferior temporal area were considered responsible for object perception, also called the 'what' pathway, while the parieto-occipital pathway was considered responsible for spatial perception, or 'where'. This concept was further refined by Goodale and Milner (1992), who instead defined the roles of the two systems by how the information processed in that stream is used: whether it was used in perception or action. The ventral stream is involved largely in the processing and identification of objects and stimuli, and was therefore considered the 'perception' pathway, while the dorsal stream is involved in tasks related to movement and action, and was thereby considered the 'action' pathway, (Ferreira et al., 1998; Goodale, 2001; Goodale & Westwood, 2004; Goodale et al., 2005; Milner & Goodale, 2008; Goodale, 2011).

Beside lesion and behavioural studies, the dorsal/ventral stream distinction is also supported by neuroanatomical evidence, where the patterns of connectivity among brain structures in the visual cortex have been found to exhibit significant level of segregation within the two pathways (Baizer et al., 1991; Kravitz et al. 2011; Kravitz et al. 2013). Segregation of visual functions is reflected in the response properties of cortical neurons (Livingstone & Hubel, 1988). In ventral stream areas, such as V4 and IT, neurons have been reported to be selective to features such as colour, complex shapes, and object identity (Desimone et al., 1984; Desimone & Schein, 1987; Gallant et al., 1993; Kobatake & Tanaka, 1994; Pasupathy & Connor, 2002). Neurons in dorsal areas, on the other hand, have been shown to be selective to the direction of motion, the patterns of optic flow fields, and disparity for stereopsis (Albright, 1984; Duffy & Wurtz, 1991; Colby et al., 1993). In particular, the large population of direction selective neurons found in the middle temporal area (MT) has made area MT the most prototypical example of dorsal stream functions. As such, the neural circuit involving area MT will be reviewed in detail in the following section.

It should be noted that despite the fact that the "two-stream hypothesis" has provided an insightful framework for understanding cortical organisation, there are significant cross-talks among dorsal stream areas and ventral stream areas. The extent to which the two-stream model accurately characterise the overall architecture of the visual cortex remains a debated issue (de Haan & Cowey, 2011).

1.2 The Middle Temporal (MT) Area

The middle temporal area (MT) is a heavily myelinated, extrastriate area that forms one of the main inputs into the dorsal processing stream (Born & Bradley, 2005), and has been identified in a number of different primates including the marmoset, the owl monkey, the macaque and humans (Allman & Kaas, 1971; Dubner & Zeki, 1971; Zeki, 1974; Rosa & Elston, 1998; Dumoulin et al., 2000; Dukelow et al., 2001). Beside its distinctive histological feature, the area MT is also unique physiologically in its simple visuotopic map (the only "first-order" map found in the visual cortex besides V1) and its large population of direction selective neurons. Studies investigating the role of MT have found it to be integral to the perception of motion: lesions of MT in macaque monkeys have shown impaired motion detection (Newsome et al., 1985; Newsome & Pare, 1988). In addition, there have been some cases where lesions in humans have also been found to lead to an inability to perceive motion (Zeki, 1991).

1.2.1 The connection patterns of MT

As a central node of the motion-processing pathway, the area MT is heavily connected to both cortical and subcortical areas (Macaque: Maunsell & van Essen, 1983a; Marmoset: Spatz, 1977; Palmer & Rosa, 2006). A simplified circuit diagram illustrating only the major pathways leading up from the retina to area MT is shown in Figure 1.1. MT receives corticocortical inputs from the primary visual area (V1), as well as inputs from V2 and V3. However, one of the things that makes MT atypical is that its activities are not entirely

dependent on input from the V1. After V1 inactivation, many extrastriate areas, such as V2 (Schiller & Malpeli, 1977; Girard & Bullier, 1989) and V4 (Girard et al., 1991), are silenced; however MT neurons retain some neuronal responses, albeit at a reduced rate (Rodman et al., 1989; Girard et al., 1992; Yu et al., 2013). This is believed to be due to direct subcortical inputs from the thalamus (Sincich et al., 2004, Bourne, 2010).

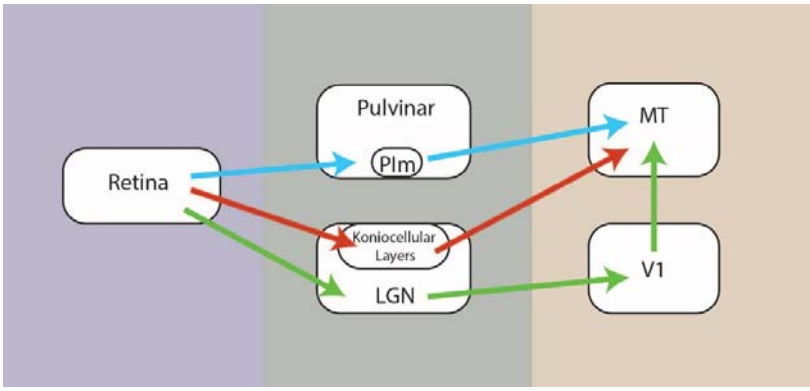

Figure 1.1. Simplified diagram illustrating the major visual pathways leading to the middle temporal (MT) area. The most prominent pathway travels from the retina through the lateral geniculate nucleus (LGN) to the primary visual cortex (V1) before reaching MT (green). In addition there are two disynaptic pathways that bypass V1 and terminate in MT: a pathway through the koniocellular layers of the LGN (red) and the medial protion of the inferior pulvinar (PIm; blue).
Adapted from Bourne, 2010.

MT acts as an input for many of the downstream areas in the dorsal pathway, with feedforward connections to areas that are considered higher-order motion processing areas (Maunsell & Van Essen, 1983a; Ungerleider & Desimone, 1986; Krubitzer & Kaas, 1990; Kaas & Morel, 1993; Rosa et al., 1993; Palmer & Rosa, 2006). In the marmoset monkey, these areas include FST (fundus of the superior temporal sulcus area), MST (the medial superior temporal area), and MTc (the crescent of the middle temporal area). There are also projections to areas in the parietal lobes, such as area VIP (ventral intraparietal) and LIP (lateral intraparietal), and some areas within the frontal lobe, including the FEF (frontal eye

fields), and adjacent granular frontal fields (Ninomiya et al., 2012). These areas have been found to be involved in many different aspects of spatial perception (processing of motion and stereopsis cues; Rosa & Elston, 1998; Tsao et al., 2003) and action (ocular tracking, saccades smooth pursuit; Bremmer et al., 2001; Goldberg et al., 2006; Ilg, 2008).

1.2.2 Physiological characteristics of MT

Within MT, there is a topographical representation of space. In other words, there is a systematic relationship between locations in the cortex and the location of the receptive fields in visual space. The visuotopic map of the marmoset MT is shown in Figure 1.2.

Receptive field sizes within MT increase along a centro-peripheral gradient such that peripheral receptive fields are larger than those located in the fovea (Albright & Desimone, 1987; Maunsell & Van Essen, 1987; Fiorani et al., 1989; Rosa & Elston, 1998). The size of the receptive fields are also larger than areas that send inputs to MT, such as areas V1 and V2 (approximately 10x and 5x respectively), when compared at similar eccentricities (Gattass & Gross, 1981; Felleman & Kaas, 1984; Albright & Desimone, 1987; Maunsell & Van Essen, 1987; Rosa & Elston, 1998).

In addition to its characteristic topographic map and receptive field size, MT also has a set of distinct, consistent functional properties. The most commonly known property is that the vast majority of neurons in the area MT, over 85%, are direction selective; meaning that the optimal response occurs in a particular direction; with a reduced, and sometimes repressed, response at 180^{o} from the preferred direction (Van Essen et al., 1981; Albright, 1984; Movshon & Newsome, 1996; Rosa et al., 2000). An example of a direction tuned response can be seen in Figure 1.3. The large population of direction selective neuron in MT is in contrast to V1, whose population of direction selective neuron has been estimated to be about 20~30% (De Valois et al., 1982; Orban et al., 1986; Sengpiel et al., 1996; Gur et al., 2005; Forte et al., 2005; Yu & Rosa, 2014).

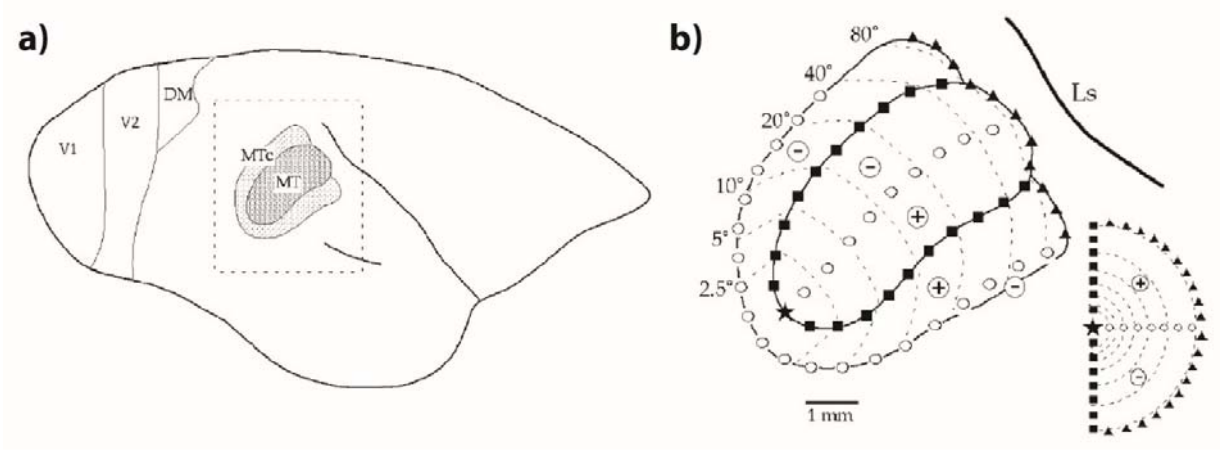

Figure 1.2. a) Lateral view of the right hemisphere of a marmoset. MT is illustrated in dark grey, and MTc is shown in in light grey. Borders of V1, V2 and DM are also indicated. The region within the dashed rectangle is expanded in b) the topographic map of the marmoset MT. The representation of the fovea is indicated by a star, while squares show the vertical meridian, and circles represent the horizontal meridian. The lower quadrant is indicated by a '-' sign, while the upper quadrant is indicated by a '+' sign, and isoeccentricity lines are shown using thin dotted lines. Figure from Rosa & Elston, 1998.

The importance of direction-tuning in the function of area MT was reflected in the columnar organisation of neurons found in a number of primate species. Rather than the orientation columns seen in the primary visual area, columns were found to be organised according to preferred direction of motion (Albright et al., 1984; Malonek et al., 1994; Diogo et al., 2003).

The role of MT in motion is firmly supported by the electrophysiological studies into neuronal properties. Yet despite this well studied set of characteristics, what remains to be determined is whether the area MT is also the location for a dense population of neurons tuned to motion speed. I will discuss this problem further in Section 1.4.

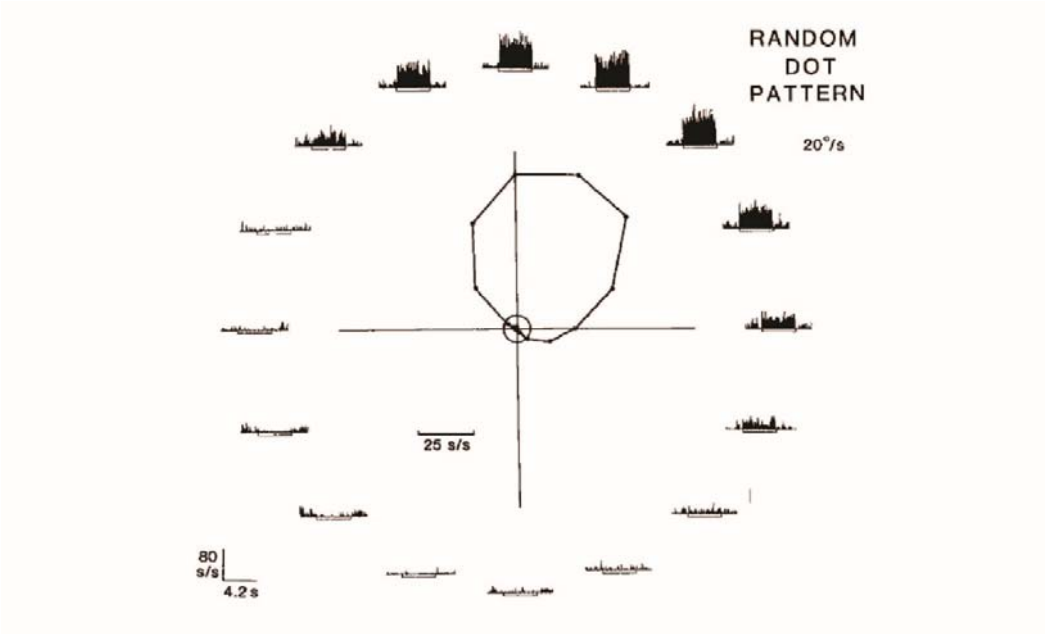

Figure 1.3. Example of the direction tuning response of a typical MT neuron. Individual histograms represent summed responses summed over 5 trials to each of 16 directions of motion of a random-dot pattern moving at 2 degrees per second. The line beneath each histogram indicates the period of time during which the stimulus was moving through the receptive field. In the centre, the response for each direction is plotted on a polar graph. The radial axis represents response (measured as spikes per second), the polar axis represents direction of stimulus motion, and the small circle represents spontaneous activity. Figure from Albright 1984.

1.3 V1 inputs to MT

One of the most distinct characteristics of the primary visual area is that it receives most of the input from the LGN, and can therefore be considered the first point of contact in the visual cortex. Because of this, the primary visual area has been well studied and has been found to have an integral role in the processing of visual information in the cortex. V1 has a clear topographical map with neurons that are characterised by small receptive fields. Populations of neurons within V1 are tuned to a variety of different visual properties including orientation, spatial frequency and colour.

In particular, V1 forms one of the main inputs into the area MT (Born & Bradley, 2005) with studies showing that V1-MT connections are highly specialised. V1 neurons that

monosynaptically project to MT have axons that are up to three times as large as other corticocortical neurons (Rockland, 1989), and their terminals are both larger and more complex than those from V2 – with connections often forming multiple synapses on a single MT neuron (Rockland, 1989; Rockland, 1995; Anderson et al., 1998; Anderson & Martin, 2002).

The response characteristics of V1 neurons that project directly to MT were studied by Movshon and Newsome (1996). These neurons were identified by applying electrical pulses to MT (antidromic stimulation) while recording in V1. The projection neurons were found to be responsive to a broad range of spatial and temporal frequencies, and were also sensitive to very low stimulus contrasts. All of the tested neurons were also found to be directionally selective. This property suggests that the direction selectivity of MT neurons are largely inherited from V1 and is supported by studies that found the average direction selectivity of MT-projecting V1 neurons to be indistinguishable from that of MT neurons.

Despite the inheritance of direction selectivity from area V1, there is evidence that MT neurons perform higher-level motion processing. One example of this is the ability for neurons to integrate multiple direction cues into a coherent interpretation of motion direction (the so-called "aperture problem", see Section 1.6.2). The MT-projecting V1 neurons identified by Movshon and Newsome (1996) were found to be unable to resolve local ambiguity in motion direction, suggesting that the ability to solve the aperture problem is a property of MT neurons, not derived from V1. Chapter 4 of my thesis compares the response characteristics of V1 and MT neurons, in order to examine if the ability for MT neurons to integrate motion signals in time is derived from V1.

1.4. The coding of motion speed in the visual cortex

In this section, I will lay out the theoretical framework of my main research. I will first explain the principles behind characterising visual neurons in the frequency (Fourier) space. In particular, emphasis is on the dependency between spatial frequency (SF) tuning and temporal frequency (TF) tuning. After reviewing the literature on the empirical measurement of SF/TF dependency, I will explain further why the dependency is at the heart of the computational question of coding motion speed in the visual cortex. I will review the current state of knowledge on the issue, and then explain the experimental design used in my thesis.

1.4.1 Characterising neuronal selectivity in the frequency domain

Fourier Theory states that an arbitrary movie, defined as a temporal sequence of 2-dimensional images can be decomposed into the linear sum of a number of components, in the form of drifting sinusoidal waves (that is, 2D sinusoidal patterns that modulate its spatial phases in time). Drifting gratings are one of the most widely used stimuli in visual neuroscience, because according to Fourier theory, the responses of a visual neuron to any arbitrary stimulus can be predicted by measuring its responses to the Fourier components; as long as the neuron acts like a linear filter. Although the linear assumption is rarely true in the visual cortex, in early visual areas, it is useful to study visual neurons' responses to Fourier components, and to characterise deviations from the linearity assumption. This method has been used successfully in the retina, the LGN, and several cortical areas, including V1 and MT (for a comprehensive review, see Shapley & Lennie, 1985).

In the visual cortex, the procedure for characterising neuronal response in the Fourier space typically employs the following procedure: First, measure the preferred direction of the neuron, using a drifting grating of fixed SF and TF. Secondly, measure SF tuning at the preferred direction, and then lastly measure TF tuning at the preferred direction and the optimal SF. In the marmoset monkey, this procedure was used to measure the SF tuning

curves and the TF tuning curves of MT neurons by Lui et al (2007a). In this study the median value for the optimal spatial frequency was 0.18 c/^{o} for neurons with receptive fields in the near peripheral visual field; the median value for the optimal temporal frequency was found to be 2.75Hz (see also Solomon et al., 2011). In a similar study in macaques, preferred values were found to be slightly higher, with the largest responses occurring between 0.5 and 1c/^{o} for spatial frequency and from 8 to 32Hz for temporal frequency (Priebe et al., 2003; Lui & Rosa, 2015).

Spatial and temporal frequency tuning curves were also measured in the V1 of marmoset monkeys (Yu et al., 2010). This study found a higher median for optimal spatial frequency of 0.48c/^{o} for neurons which had receptive fields centred in the near periphery, at similar eccentricities to that of Lui et al (2007a), and my own research. The optimal temporal frequency was comparable across the visual field (median value 3Hz).

The practice of testing SF and TF tuning separately was introduced to reduce the long experimental time needed to test all possible combinations of SF and TF. The procedure gives a complete measurement of a neuron's spatiotemporal characteristics only if its SF-tuning and TF-tuning are *separable*, meaning that the SF-tuning curve of a neuron is not dependent on the TF of the drift grating used to make the measurement (and vice versa). If such a condition is met, a neuron's complete spatiotemporal "tuning surface" can be inferred by its SF-tuning curve and its TF-tuning curve by calculating the product of the two curves. Formally, the relationship can be written as

$$M_0(sf, tf) = b + c \cdot M_s(sf) * M_{tf}(tf) \tag{1.1}$$

where

$$M_s(sf) = \exp\left[\frac{-(sf-sf_0)^2}{2\left(\sigma_{sf}-\zeta_{sf}(sf-sf_0)\right)^2}\right] - \exp\left[\frac{-1}{\zeta_{sf}^2}\right] \tag{1.2}$$

$$M_t(tf) = \exp\left[\frac{-(tf-tf_0)^2}{2\left(\sigma_{tf}-\zeta_{tf}(tf-tf_0)\right)^2}\right] - \exp\left[\frac{-1}{\zeta_{tf}^2}\right] \quad (1.3)$$

Both the SF and TF tuning curves (M_s and M_t) are modelled as skewed-Gaussian distributions in logarithmic scale (Log_2 transformed), where σ indicates the bandwidth of the neuron, ζ indicates the skewness of the response; sf_0 and tf_0 represent the optimal SF and TF after Log_2 transformation. The value of *b* represents the spontaneous activity of the neuron, while *c* provides a scaling factor. This process is illustrated in Figure 1.4, where Figure 1.4c shows the response profile for M_0.

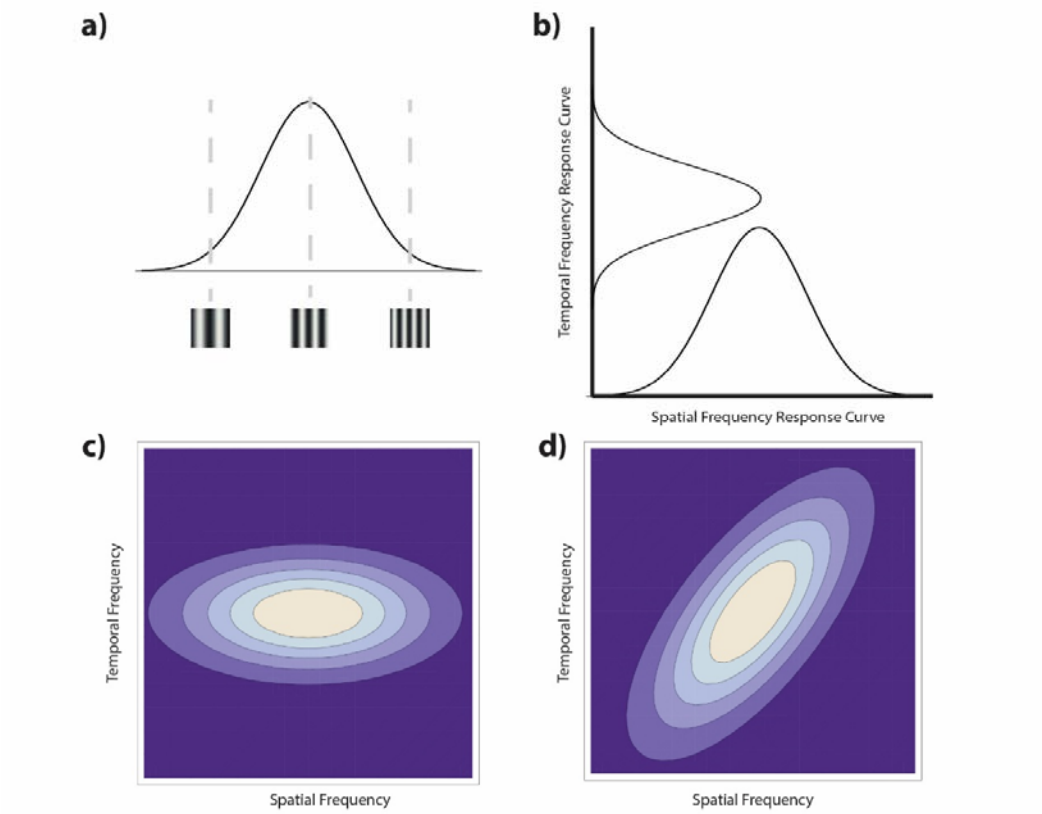

Figure 1.4. a) A theoretical tuning curve of a neuron to spatial frequency as it varies along the x-axis with firing rate shown along the y-axis. b) The response characteristics of a neuron is very often quantified by the SF and the TF tuning curve, although it should be measured as a tuning surface. c) Contour plot illustrating the SF-TF tuning surface of a spatiotemporally separable neuron. d) Contour plot illustrating the SF-TF tuning surface of a speed-tuned neuron. The x and the y axis are in log-log scale

The separability assumption is very often not empirically tested. However, in macaque V1, Foster et al. (1985) explicitly checked the assumption by measuring V1 neurons' responses to combinations of SF and TF, and found that the neuron's preference for spatial frequency was essentially independent of the test temporal frequency. The issue was revisited by Priebe et al. (2006), who reported a broad distribution of separability in direction selective neurons in the macaque V1. In the marmoset V1, Yu et al (2010) found that that SF-tuning and TF-tuning were separable, except that there was a slight increase of SF-TF dependency in the far peripheral visual field.

The issue of SF/TF separability is more critical in area MT, because it is at the heart of the problem of speed coding. This problem is described in the next section.

1.4.2. The computational problem of speed tuning in the Fourier space

The Fourier space provides an insightful framework for studying the neural processing of moving signals. The full Fourier theory for motion processing deals with images moving at arbitrary directions on the 2D plane ("2D motion"), and involves computation in 3D space - 2 dimensions for spatial frequency and 1 dimension for temporal frequency (Watson & Ahumada, 1983; Watson & Ahumada, 1985; Simoncelli & Heeger, 1998). For the purpose of my thesis, I will concentrate on a simplified version of the theory, which only addresses motion at the preferred direction of a neuron ("1D motion"). Motion at non-preferred directions is outside the scope of the thesis, and needs to be further investigated with a different research paradigm (see Nishimoto & Gallant, 2011, for an example). In the context of my research, motion processing is a computational problem in the 2D Fourier space (1 dimension for spatial frequency and 1 dimension for temporal frequency).

The Fourier representation of a spatially restricted pattern moving at constant speeds at a fixed direction (such as the preferred direction of a neuron) is illustrated in Figure 1.5a. In this figure, the x-axis represents SF, the y-axis represents TF, and the lines represent the

concentration of Fourier energy of different speeds of motion. Lines with lower slopes correspond to motion at slower speeds. The same SF-TF relationship is re-plotted in double log scale in Figure 1.5b. In this plot, motions of different speeds have the same slope (slope=1), but different intercepts. I will use the double-log scale for plotting the 2D Fourier space in the rest of the thesis.

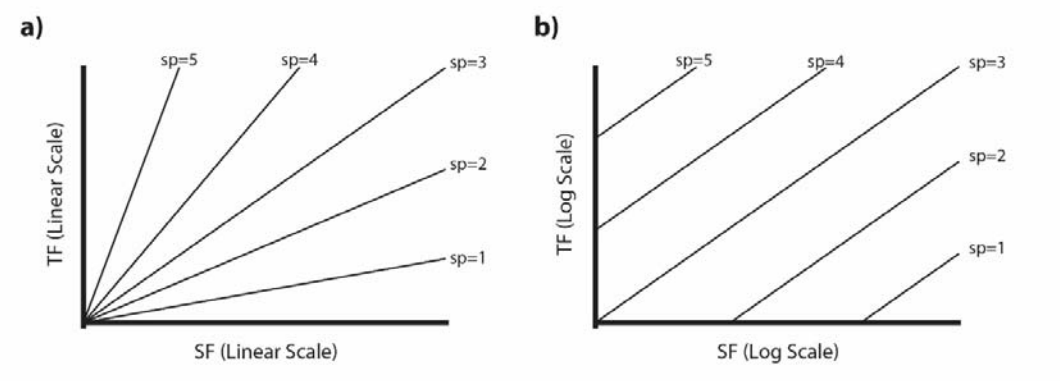

Figure 1.5. Motion of different speeds, represented in the Fourier space. In a), the SF and the TF are in linear scale. In b), they are in logarithmic scale.

Figure 1.5b shows that for a neuron to be tuned to the speed of a moving pattern, its selectivity to the TF of the stimulus must be dependent on the SF of the stimulus. Furthermore, the dependency must have the slope of 1 in double log scale. Neurons with separable SF-TF tuning, such as the vast majority of V1 neurons discussed in the previous section, encode the TF of a stimulus independently of its SF content, and therefore cannot be speed-tuned. In a condition where the SF of the stimulus is fixed, the response rate of such a neuron is correlated with the speed of a moving stimulus. However, in a condition where the SF varies, the response rate is *ambiguous* with regard to speed - in the sense that the same response rate can have multiple interpretations. To see this, consider a TF-tuned neuron responding to two moving patterns: one with a SF distribution concentrating on 1c/^{o}, moving at the speed of 2^{o}/s, the other with the same SF distribution but shifted to 2c/^{o}, moving the speed of 1^{o}/s. Since speed is the ratio between TF and SF (Fig. 1.5), both stimuli have the

same TF (2 Hz). A neuron with separable SF-TF tuning would therefore respond to both stimuli with the same response rate; the same response rate can be caused by a wide variety of speeds, and therefore is ambiguous.

In a complex environment where SF content can vary significantly across the visual field, this ambiguity can be a significant problem. Consider a zebra running across the visual field (Fig. 1.6). To predict the location of the zebra at a future time point, it is essential that the entire zebra is represented as an object moving at a coherent speed. But given the variation in SF across the body of the zebra, the responses of TF-tuned neurons would vary at different locations on the body, and cannot signal the coherent moving speed of the zebra.

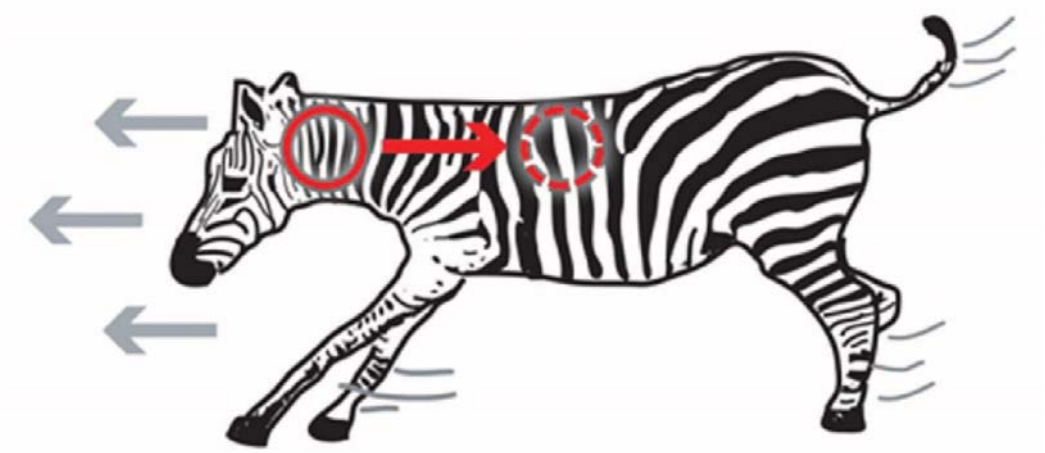

Figure 1.6. On different parts of the zebra, the stripes vary in width and density, creating different spatial frequencies. On the neck of the zebra, indicated by the solid circle, the stripes are finer and more densely packed, representing a higher spatial frequency. While the stripes along the back are thicker, showing a lower spatial frequency. For speed-tuning accuracy, neurons must be invariant to spatial frequency, otherwise different speeds would be indicated at different parts of the zebra. Figure from Moore, 2006.

1.4.3. Quantifying speed tuning

Early studies of speed-tuning in the visual cortex measured the responses of a neuron to a bar moving at the preferred direction of the neuron, at various speeds (Orban et al., 1981;

Maunsell & van Essen, 1983b; Felleman & Kaas, 1984; Orban et al., 1986; Rodman & Albright, 1987; Lagae et al., 1993; Cheng et al., 1994). If the responses are only significant in certain intervals of speed, the neuron is said to be speed-tuned. The method, however, is flawed, because the broad distribution of SF in the bar stimulus makes it impossible to determine if the neuron is tuned to the speed of the stimulus, or to the TF of the stimulus.

This led to the development of a more sophisticated method using stimuli with narrow-band SF contents – namely drifting gratings. Perrone and Thiele (2001) introduced the concept of measuring the spatiotemporal tuning surface using drifting gratings whose SF and TF are independently varied. A popular method for fitting the tuning surface was introduced by Priebe et al (2003) who used a Q parameter to quantify the dependency between SF and TF tuning. A slightly modified formulation of this method, which has been used to study speed tuning in the marmoset monkey (Lui et al., 2007a; Yu et al., 2010) is as follows:

$$M_{tf}(sf, tf) = \exp\left[\frac{-(tf - tf_{sf})^2}{2\left(\sigma_{tf} - \zeta_{tf}(tf - tf0)\right)^2}\right] - \exp\left[\frac{-1}{\zeta_{tf}^2}\right] \quad (1.4)$$

$$tf_{sf} = Q \cdot (sf - sf0) + tf0 \quad (1.5)$$

$$M_1 = b + c \cdot M_{sf}(sf) \cdot M_{tf}(sf/tf) \quad (1.6)$$

In this equation, the Q parameter represents the rate at which the optimal TF varies with the SF of the stimulus (in double log scale). Values of Q fall along a continuum where a Q of 0 indicates no dependence between the variables and is equivalent to M_0 (Equation 1.1), and a Q of 1 indicates an idealistically speed-tuned, neuron (Fig. 1.4d). Values that fall between indicate that while the optimal temporal frequency increases with spatial frequency, it does so at a level lower than that expected of a perfectly speed-tuned neuron. Similarly, for values of Q above 1, the optimal temporal frequency is increasing at a rate higher than that expected of a speed-tuned neuron.

An issue involved in the determination of SF-TF separability using this formulation is that an arbitrarily chosen threshold for Q is required. To void imposing an arbitrary criterion on the "how small should Q be for a neuron to be speed-tuned" question, Yu et al. (2010) introduced a Bayesian procedure, which fitted the same tuning surface to two competing models, one with the Q parameter (Equation 1.4) and one without (Equation 1.1), and used the Bayesian Information Criterion to select one of the two models which was more likely to provide better explanation of the data.

Data analysis in my thesis was performed under the same framework, except that the tuning surface had to be formulated in SF-speed coordinates instead of in SF-TF coordinates, to incorporate the new stimulus design. This will be explained in Section 2.3.1.

It should be noted that alternative frameworks for quantifying speed tuning exists. For example, Perrone & Thiele (2002) proposed the weighted intersection mechanism (WIM) model, which uses combined inputs of transient and sustained neuron types to generate the MT speed-tuning response.

1.4.4. Speed tuning in MT

Early studies investigated the responses of MT cells to broadband stimuli such as moving dot fields and single bars (Maunsell & van Essen, 1983b; Allman et al., 1985; Tanaka et al., 1986; Snowden et al., 1992; Britten et al., 1993; Cheng et al., 1994). These experiments reported that the majority of MT neurons are tuned to the speed of a moving stimulus. However, as discussed in Section 1.4.3, these experiments confounded SF and speed, and therefore do not provide strong conclusions about speed tuning.

The use of Fourier components for the study of speed tuning was first introduced by Perrone and Thiele (2001), who reported that a large population (61%) of MT neurons were SF-TF inseparable. However, later studies using the same paradigm but a more stringent criterion for

speed tuning have found smaller populations. A study in anaesthetised macaques reported that only a quarter of MT neurons responding to speed invariant of spatial frequency (Priebe et al., 2003). An additional study in awake behaving macaques found that most neurons were not tuned for a consistent speed (Miura et al., 2014), with 10% of neurons responding in a spatial frequency-invariant manner. In anaesthetised marmosets, Lui et al. (2007a) found only 10% of tested MT neurons to be classified as speed-tuned. In summary, despite MT being one of the centres of motion processing in the primate visual cortex, only a small proportion of MT neurons have been found to be speed-tuned, when quantified with drifting gratings.

1.5. Visual Stimuli for Characterising Motion Selectivity

Since human subjects can make precise judgements on the speed of drifting sinusoidal gratings in psychophysics experiments (McKee et al., 1986; Reisback & Gegenfurtner, 1999), it is somewhat puzzling as to why most studies have reported small populations of speed-tuned neurons in MT. It may be that the visual cortex does not require a large population of speed-tuned neurons in order to solve complex visual tasks involving moving objects. The small population of speed-tuned neurons therefore does not present a problem. However, since a moving pattern is jointly specified by its direction and its moving speed, it is unclear why a large population of direction selective neurons are found in MT, while only a small popoulation of speed-tuned neurons are found there. No existing theory explains this direction/speed asymmetry.

Alternatively, the population of speed-tuned neurons may be located in visual areas at higher level in the hierarchy of motion processing. An fMRI study by Lingnau et al. (2009) observed SF-invariant speed coding in human area MT and MST, suggesting that MT (at least in human) should play significant roles in processing image speed, but it is not the only cortical area that does so.

Another possibility is that the drifting gratings that have been used in previous studies are devoid of the complex features found in naturally occuring motion signals, and therefore fail to reveal the full extent of complexity of motion processing. This possibility was investigated by Priebe et al. (2003), where they studied the effect of increasing the spatial complexity of the stimulus on speed-tuning. Natural images are rich in their spatial-frequency contents, whereas the sinusoidal stimulus contains only a small range of spatial-frequencies. By superimposing two sinusoidal patterns containing different spatial frequency content, Priebe et al. (2003) found that MT neurons are more speed-tuned when they are characterised by the superimposed stimuli, in comparison to situations where they are characterised by single-component sinusoidal gratings.

The novel stimulus design used in my thesis follows the same direction, except that my research was focused on the temporal aspect of the moving signal. An important feature of naturally occurring motion is that motion signals are very often caused by objects moving in the visual field, which travel in trajectories. A trajectory will result in the consecutive activation of neurons along a path and, while this feature was present in the moving bar stimulus used in earlier studies, it is absent in drifting gratings, which only carry local motion signals. In other words, the motion stimuli used in the majority of motion research do not actually move in space.

In my thesis, I introduce a "Moving mode" stimulus, which is a static sine-wave grating located within a moving envelope. In this stimulus, speed is represented by movement of the envelope across the visual field, as it stimulates consecutive receptive fields along a path. This is in contrast to the "Standard mode" (the stimulus used by most grating studies), where speed is created through time-modulation of the sine-wave grating, while the envelope

remains static, in the same area of the receptive field throughout the duration of the presentation time.

A third mode, the Waxing/Waning mode, was used to control the temporal differences between Moving mode and the Standard mode. The three stimulus modes are described in detail in Section 2.2.3.

1.6 Computational models of motion processing

In Chapter 5 of my thesis, I used computational models to address theoretical issues raised by the experimental results in Chapter 3 and Chapter 4. In this section, I will establish the theoretical foundation of the modelling research, and explain the questions that the computation model was designed to address.

1.6.1 The motion energy model

Computational models of motion processing in the visual cortex typically begin with elementary motion detectors – units that respond to movement at one particular direction and not the opposite direction, at localised regions of the visual field. Although different formulations of elementary motion detectors have been proposed (Burr & Thompson, 2011), in my thesis I utilised the most commonly used model – the spatiotemporal-energy (STE) model proposed by Adelson and Bergen (1986). The choice of model is not a significant issue because the properties of other models (such as the Reichardt detector) are not qualitatively different from the energy model, in the context of my research.

The spatiotemporal-energy (STE) models (Fig. 1.7) account for a wide variety of motion data and are based on physiologically plausible mathematical operations (Emerson et al., 1992; Touryan et al., 2002; Rust et al., 2005; Bradley & Goyal, 2008). The first step in the STE model is a linear filtering step, where an arbitrary stimulus (that is, a movie) is convolved with a spatiotemporal filter, which is typically chosen to be a drifting Gabor filter (bandpass

filter in SF and TF in the frequency domain; see 2.2.3). The direction of the drift establishes the direction selectivity of the STE.

The second step calculates motion energy. For a drifting sinusoidal pattern, the output of a linear filter will fluctuate between positive and negative values in a sinusoidal wave – which is biologically unrealistic because a neuron cannot have a negative firing rate, and tends not to fluctuate in synchrony with the drifting grating. To solve this problem, a quadrature pair is used, meaning that the input movie is convolved with a pair of drifting Gabor filters, which have the same spatial and temporal frequency, with a 90 degree offset in their spatial phases. The response from each filter pair are then squared and summed. Since the outputs of the two filters fluctuate in sinusoidal patterns with 90 degree offset (that is, one sine wave and the other cosine wave), when they are squared and then summed, the result is a constant positive value. This behaviour reflects that of complex cells in V1, which have been shown to respond to drifting sinusoidal gratings in a phase-insensitive manner (Emerson et al, 1992).

The final step in the model is referred to as opponency, and is targeted at reducing noise in the signal. In this step, the motion energy of one direction is subtracted by the motion energy of the opposite direction. If noise in the input signal is omnidirectional, it is cancelled out.

The energy model captures several characteristics of direction selective neurons found in V1: it responds to motion in a particular direction (determined by the filter) and not to the opposite direction. Furthermore, its SF- and TF-tuning are separable, and therefore is not speed-tuned.

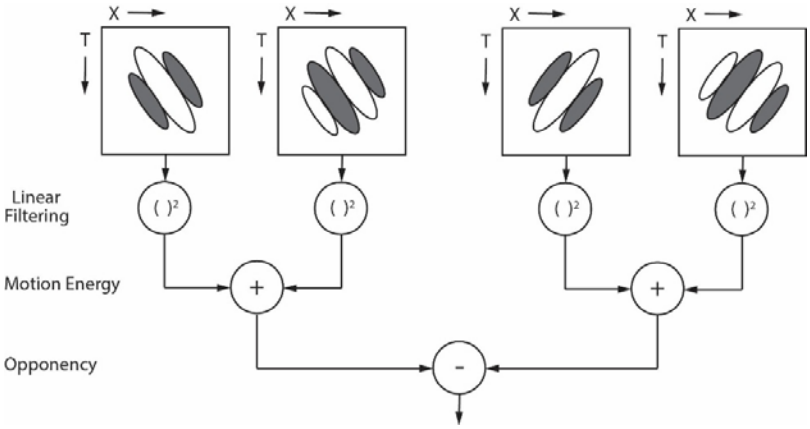

Figure 1.7. A schematic depiction of the basic operation of the spatiotemporal-spatiotemporal-energy (STE) model of elementary motion detector. See Section 1.6.1 for details.

1.6.2 The integration of local motion signals

Given that many important features of direction selective neurons in V1 are captured by the energy model, the question that naturally arises is the computation performed at the level of MT. The most prominent theory for computation at the level of MT was proposed by Simoncelli and Heeger (1998), which stipulates that MT neurons operate as integrators of local motion signals that originate from V1. Since the model was proposed as solution to the "aperture problem", the rest of this section will be devoted to the description of the problem, and Simoncelli and Heeger's solution to the problem. The relationship between the aperture problem and the problem of speed tuning will be examined in the next section.

The aperture problem is a computation problem in motion processing, which originates from the fact that a single neuron only has access to a limited region of the Fourier space, and as a consequence, cannot always encode the motion signal unambiguously. Despite its name, the aperture problem is a generic problem that can manifest itself in many different forms, which do not necessarily need to involve an aperture (Bradley & Goyal, 2008). For simplicity, I will focus on the special case that does involve an aperture. A visual neuron "sees" the world

thorough an aperture, its receptive field. Figure 1.8 illustrates a surface moving towards the right. However, when it is viewed through an aperture, the surface appears to be moving towards the upper-right. This is because the vector representing the motion direction of the surface can be decomposed into two orthogonal vectors. The vector that is perpendicular to the edge of the surface is visible to the receptive field of the neuron, but the vector running parallel to the edge is not. The neuron therefore is only able to detect the visible vector, which is not the true motion direction of the surface.

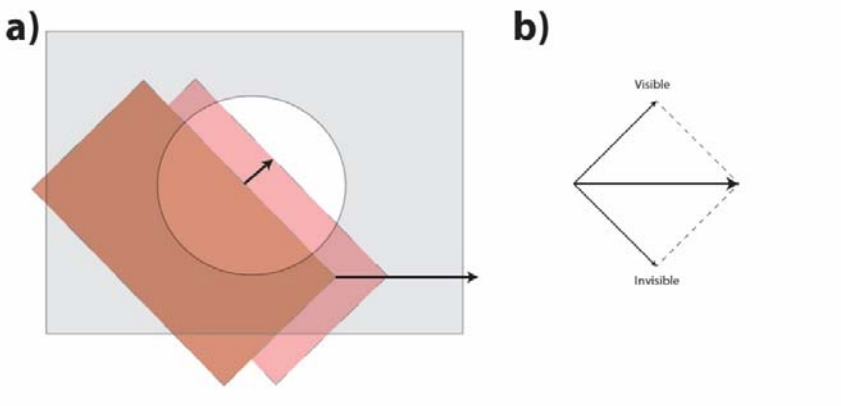

Figure 1.8. The aperture problem. a) Even though the rectangle is moving directly to the right, apparent motion as seen through the aperture (indicated by the central circle) is upwards and to the right. This is due to only one of the orthogonal vectors being visible as see through an aperture. b) shows these two orthogonal vectors (both the visible and invisible vectors) as well as the actual motion indicated by the middle arrow.

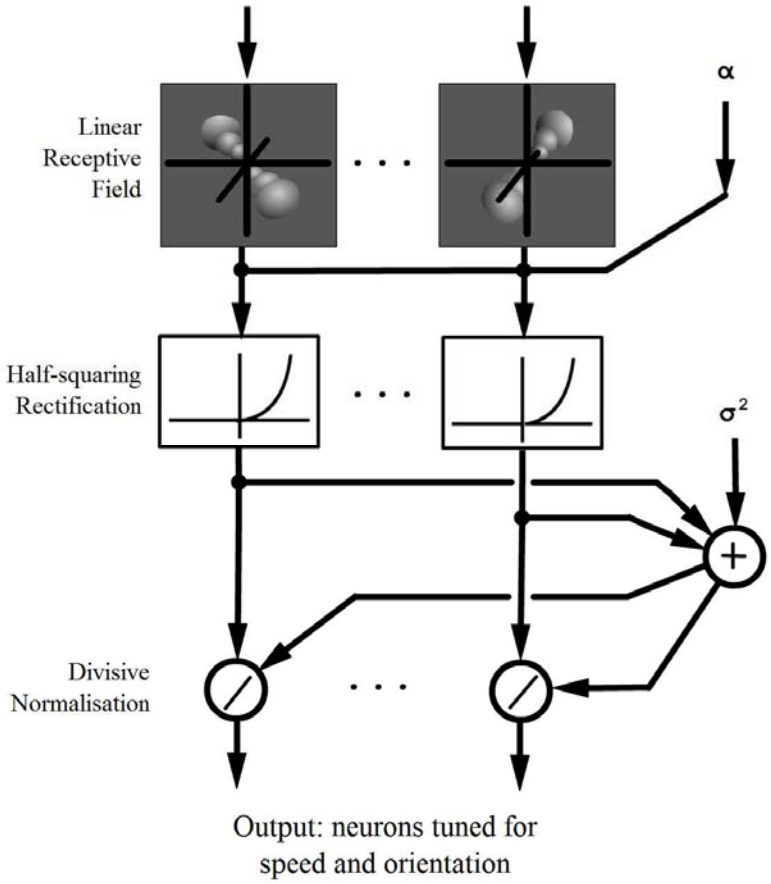

Figure 1.9. The Simoncelli and Heeger model of MT pattern cells. Each neuron computes a weighted sum of V1 complex cell afferents (each blob in the Linear Receptive Field panel represents the spatiotemporal tuning surface of a V1 neuron in the 3D Fourier space), followed by half-squaring, and normalisation. The V1 afferents are weighted to give a velocity-selective response. The divisive normalisation factor is computed as a sum of half-squared responses and a squared semi-saturation constant. Figure replicated from Simoncelli and Heeger, 1998.

The computational model of Simoncelli and Heeger (1998; Fig. 1.9) proposed that the aperture problem can be solved by integrating (that is, "pooling") the outputs of multiple motion energy detectors. The model involves three steps: V1 outputs are first summed linearly by a set of weights, which is followed by half-squaring rectification, and then divisive normalisation (Carandini et al., 1997; Carandini & Heeger, 2012). In the half-squaring rectification step, if the output is less than zero, it is set to 0. Otherwise, it is squared. In the normalisation step, the rectified outputs are divided by the sums of all output units.

The behaviour of the Simoncelli and Heeger model closely mirrors that of a class of MT neurons called "pattern cells", which are identified by measuring the responses of the neuron to plaids (that is, the summation of two drifting sinusoidal gratings) of different directions. A neuron is classified as a pattern cell if its tuning curve reflects the global motion direction of the plaid, instead of the individual gratings that make up the plaid (Movshon et al., 1985). In other words, patterns cells solve the aperture problem. In the macaque and the marmoset MT, it has been reported that about 20% of MT neurons are pattern cells (Movshon et al., 1985; Smith et al., 2005; Solomon et al., 2011).

1.6.3 The aperture problem and speed tuning

In Section 1.4.2, I introduce the distinction between 1D motion (a pattern moving at the preferred direction of a neuron) and 2D motion (a pattern moving at any arbitrary direction). The distinction is useful for explaining the relationship between the Simoncelli & Heeger model and my research. In 2D motion, the local motion signals from elementary motion detectors (such as the energy model described in Section 1.6.1) are ambiguous in both the direction and the speed of motion. Both ambiguities are solved by the Simoncelli and Heeger model, because the model is constructed to resolve the ambiguity of the motion vector, which comprises both the direction and the speed. In 1D motion (the scope of my research), there is no ambiguity in the direction of motion, since the direction of motion is always the preferred direction of a neuron. The aperture problem is therefore reduced to the problem of ambiguity in speed. In this context, the Simoncelli and Heeger model is therefore reduced to a mechanism for resolving ambiguity in speed. In both contexts, ambiguities can be resolved by integrating the outputs of elementary motion detectors.

The significance of the Simoncelli and Heeger model is that it explains the existence of "pattern cells" in MT (Section 1.6.2). A stronger prediction is that pattern cells in MT should also be speed-tuned. This prediction was explicitedly tested by Priebe et al. (2003), who

showed that pattern cells are not always speed-tuned. The result therefore calls the model into question.

1.6.4 Linear summation in motion integration

A critical feature of the Simoncelli and Heeger model is that despite the nonlinear elements in the model, at the stage where signals from the elementary detectors are integrated ("pooled"), a linear combination of the inputs is used (Rust et al., 2006; Nishimoto & Gallant, 2011). A consequence of the linear summation mechanism is that the behaviour of the neuron is entirely predictable from its responses to Fourier components (drifting gratings). A neuron that is not speed-tuned when it is tested with drifting gratings should also be not speed-tuned when it is tested with a stimulus constructed from linear combinations of drifting gratings. This prediction was explicitly tested by Preibe et al. (2003), and proven to be inaccurate (Section 1.5). The result suggests a certain degree of nonlinear summation in the operation of MT neurons, and indicates that the Simoncelli and Heeger model does not capture the entire complexity of motion processing in MT.

Since the Moving mode stimulus used in my thesis can be decomposed into a linear summation of Fourier components (drifting components), the experiments can be understood as a test for the linearity assumption in the same vein as Priebe et al. (2003). My approach is different from Priebe et al. (2003) in two aspects:

1. In Priebe et al. (2003), two Fourier components are added to form a new stimulus. In my research, the Moving mode stimulus is the summation of many (in fact, infinitely many in an idealised situation) Fourier components

2. In Priebe et al. (2003), the summation of the two Fourier component does not result in a displacement of the envelope. In my stimulus, the parameters of the components were chosen specifically to result in a motion trajectory.

To ensure that the results described in Chapter 3 are not the expected behaviour of a Simoncelli and Heeger-type model, I constructed such a model, fitted to the SF-TF tuning surface measured with the Standard mode stimulus, and used it predict the behavior of the model when it is stimulated by the Moving mode stimulus. The results confirmed that despite the nonlinear components included in the model, the results were not expected. The result provides further evidence for nonlinear summation in MT.

1.7 The marmoset as a model animal in neuroscience research

The common marmoset (*Callithrix jacchus*) is a small new-world monkey with a fast maturation rate that has been found to be able to breed more rapidly in captivity than other primates (Rylands, 1993). This is desirable in a primate model as the animals mature within 18 months, decreasing the strain of breeding and housing until maturation. Also, because of its small size it is easy to maintain family groups in captivity, which in turn reduces potential boredom and possibility of neuroses sometimes seen in macaques and baboons caged singly (Hearn, 1983). The marmoset brain is lissencephalic, with a surface area more than 10x smaller than that of the macaque (Chaplin et al., 2013), allowing for easy navigation of cortical areas during electrophysiological recordings. In addition, the availability of multiple transgenic lines makes the marmoset a highly important model in biomedical research (Sasaki et al., 2009).

The marmoset is fast becoming one of the most popular primate models in neuroscience research. Advanced research methods have been developed for the study of the marmoset, which include functional MRI (Hung et al., 2015), behavioural training (Yamazaki et al., 2011; Mitchell et al., 2014), array electrophysiological recording (Zavitz et al., 2016), two-photon imaging (Sadakane et al., 2015), and optogenetics (MacDougall et al., 2016). The visual system of the marmoset is very well-studied (Mitchell & Leopold, 2015; Solomon & Rosa, 2015), with most of the visual cortex mapped in detail (Rosa & Tweedale, 2005;

Paxinos et al., 2012). Visual areas V1 and MT in the marmoset have been studied in great detail (Fritsches & Rosa, 1996; Rosa et al., 2000; Bourne et al., 2002; Lui et al., 2007b; Solomon et al., 2011), and their characteristics have been shown to mirror closely those that have been reported in the macaque monkey (Solomon et al., 2011; Lui & Rosa, 2015) – the other dominant primate model for the study of the visual cortex.

1.8 Aims of Thesis

Aim #1: To determine if patterns that move in predictable trajectories increase the extent to which a single neuron in area MT unambiguously encodes the speed of motion.

Aim #2: To examine if the effect studied in Aim #1 is also present in V1.

Aim #3: To examine if the context-dependent speed-coding studied in Aim #1 can be explained by the classic model of MT neurons.

2 General Methods

Experimental protocol for this project involved the surgical preparation of the animal, followed by an extended period of electrophysiological recordings (typically lasting 3 days). Experiments were ended through the humane sacrifice of the animal by Pentobarbitone sodium injection, and the brain was extracted, sectioned and stained, to confirm that the recordings were obtained in the intended cortical areas (V1 and MT). Data were analysed with programs written in *Mathematica*.

2.1 Animal Preparation

Single-unit recordings were obtained from the primary visual cortex (V1) and area MT in 5 adult New World monkeys (*Callithrix jacchus*, the common marmoset). Apart from one animal in which both areas were studied, electrophysiological recordings were performed in either MT (n = 2) or V1 (n=2). Experiments were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes, and all procedures were approved by the Monash University Animal Ethics Experimentation Committee.

Preparation of the animal for *in vivo* recordings of neuronal responses involves the use of several surgical stages, based on the protocol described in Bourne and Rosa (2003). One amendment to this published protocol was made in the use of Alfaxan in place of ketamine/xylazine for the initial induction of anaesthesia.

2.1.1 Medication

The pre-anaesthetic medication comprised 3mg kg^{-1} diazepam (Pamlin, CEVA) and 0.2mg kg^{-1} atropine (Atrosite, Ilium) which were both administered by intramuscular injection, given half an hour before induction of anaesthesia. Diazepam was used to induce a mild level of sedation. Atropine was used to reduce bronchial and salivary secretion level, and to minimize cardiac rhythm disturbance.

The induction of anaesthesia used Alfaxan (Jurox, 10mg kg^{-1}), which kept the animal anaesthetised for the duration of the surgical procedures (under two hours). Other injections administered during the surgery were dexamethosone (Dexason, Ilium, 0.3mg kg^{-1}), injected intramuscularly as a measure to prevent cerebral oedema, and penicillin (Norocillin, Norbrook, 25mg kg^{-1}), also injected in the muscle, used as an antibiotic. A long-lasting local anaesthetic (Xylocaine/Adrenaline, 0.5%/1:200,000, AstraZeneca) was injected locally at the areas of surgical incision, namely the thigh, neck and top of the cranium.

Once the surgical procedures (described in detail below) were completed, we initiated injection of a maintenance solution intravenously, combined with artificial ventilation. The maintenance solution was constantly administered at a level of 5ml hr^{-1} for the first half hour, before being reduced to a lower rate of 2ml hr^{-1}. The solution included sufentanil citrate (an opioid anaesthetic, injected at a rate of 2.66 $\mu g.kg^{-1}.h^{-1}$), pancuronium bromide (a nicotinic blocker, which resulted in relaxation of the eye musculature, injected at a rate of 1.44 $mg.kg^{-1}.h^{-1}$), Dexapent (Ilium) a potent anti-inflammatory, injected at a rate of 7.220 $\mu g.kg^{-1}.h^{-1}$), and Xylazil (a sedative, analgesic and muscle relaxant, injected at a rate of .823 $mg.kg^{-1}.h^{-1}$). These drugs were injected in a solution of balanced salts and nutrients (Synthamin-13, Hartmann's solution, and a .18% sodium chloride/4% glucose solution in a proportion of 12:30:25 in volume). This solution results in a sustained level of deep anaesthesia for the duration of the experiment, particularly when combined with artificial ventilation with a gaseous mixture of N_2O and O_2 (7:3). The rate of infusion was adjusted over the course of the experiment (typically, 2 $ml.h^{-1}$) to maintain an optimal heart rate of between 150 to 250 beats per minute.

At the conclusion of the experiment the animal was humanely killed through the injection of Pentobarbitone sodium at 100mg kg^{-1}.

Throughout the experiment, physiological parameters were monitored to ensure that the level of anaesthesia was maintained. These included heartrate, blood oxygenation level, body temperature, and the spontaneous activity of the brain.

2.1.2 Surgery

After the animal was anesthetised, two surgeries were performed. A tracheotomy allowing for artificial ventilation throughout the experiment, and a femoral cannulation which allowed for administration of the intravenous maintenance infusion.

2.1.2.1 Tracheotomy

The first procedure performed was the tracheotomy. Using a scalpel, an incision was made along the throat, and blunt dissectioning techniques were used to reveal the trachea. Two pieces of surgical thread were fed underneath the trachea in preparation for tying off the tracheal tube, and to ensure tracheal tube had been fully revealed. A small cut was made between rings of the upper tracheal cartilage to allow insertion of a tracheal tube. Surgical thread was then tightened around the tube both above and below the incision line, and fast-drying cyanoacrylate glue (Supa Glue, Selleys, Australia) was applied to the top of the knots to ensure they remained in the appropriate area, minimizing the possibility of the tracheal tube being dislodged. The site was also closed with the same glue. Standard commercial tracheal tubes (inner diameter 2.5 mm) were used but were trimmed back to 54 mm length to minimise dead air space during ventilation.

2.1.2.2 Cannulation

The next step involved cannulation of the femoral vein, and was done through a similar method. A scalpel was used to make a longitudinal incision along the inner thigh, and fine scissors used for blunt dissection in order to reveal the vein. After a 1 cm long segment of the vein was isolated from the femoral artery and from all connective tissue, surgical thread was used to tie off the distal end, preventing centripetal blood flow. A microincision was then made using iridectomy scissors, and the cannula inserted towards the heart, for approximately 40mm. The cannula was filled with saline before insertion and attached to a syringe of saline. The syringe was pulled back until blood was seen in the catheter in order to confirm insertion in a vein. The segment of the vein proximal to the microincision was tied off, securing the cannula in place, and the knot secured with cyanoacrylate glue. The surgical site was also closed with cyanoacrylate glue.

2.1.2.3 Craniotomy

After the cannulation was performed the marmoset was moved to a stereotaxic frame, a diagram of which can be seen in Figure 2.1. The head was secured via a combination of ear bars (coated with xylocaine jelly), a palatine bar, and two lower orbital bars. Intravenous injection of normal saline was initiated at this point, to avoid blockage of the vein cannula, and sensors for temperature and blood pressure/ oxygenation were attached to the animal.

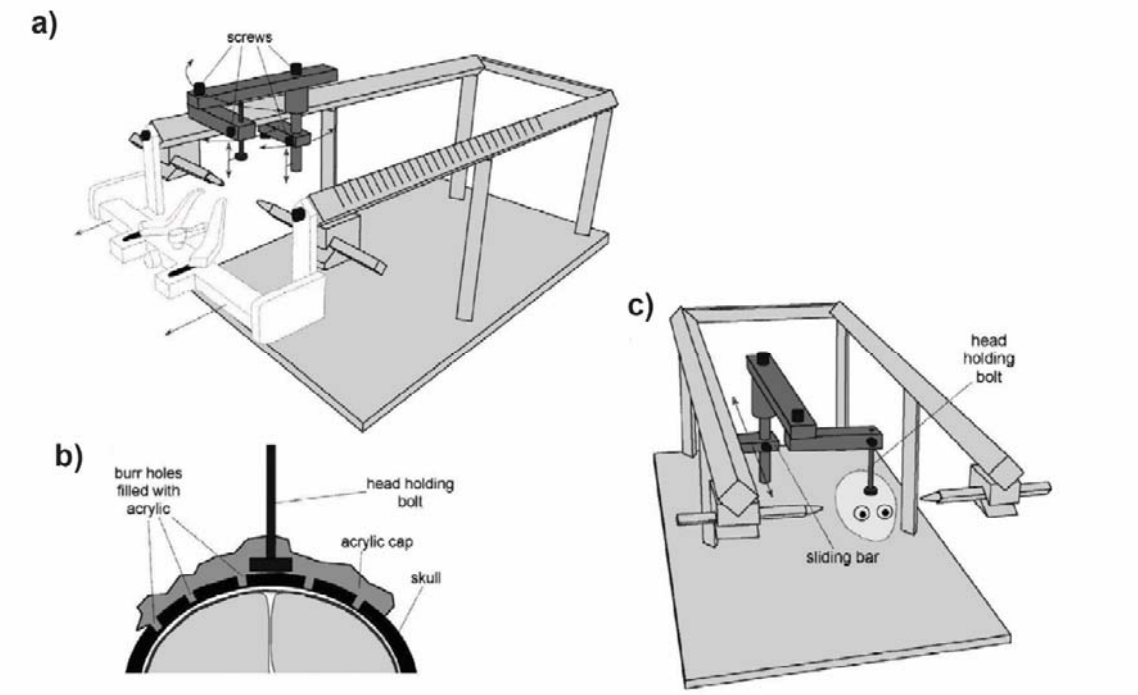

Figure 2.1. Method for holding the marmoset head in position for duration of the experiment. a) shows the complete stereotactic frame. The orbital and palatal bars are shown in white and are used during the craniotomy and before attachment of a bolt to the skull. After attachment of the head holding bolt these are removed to allow for unobstructed view and the head is supported by the bars marked in dark gray. b) shows attachment of the bolt to the skull; several burr holes are drilled into the skull to allow for extra purchase of acrylic glue, which is then used to attach the bolt to the skull. c) shows the final set up, with the head holding bolt and ear bars now responsible for holding the head in place.

Images from Bourne and Rosa (2003)

Once these steps were completed, a craniotomy was performed to make the area of recording accessible to the electrodes. A small opening was made over area MT in the right hemisphere,

approximately between anteroposterior levels -1 and +6 (mm from the ear bars). For V1 experiments both hemispheres were exposed by making an opening over the posterior portion of the skull. The dura mater was also removed over the craniotomy areas, to avoid damage to the tip of the electrodes, using a pair of fine forceps and dura scissors.

2.1.3 Focusing and Care of the Eyes

Atropine (Atropt 1%) was used to dilate the pupil of the eye (mydriasis), and phenylephrine paralysed the muscles responsible for accommodation of the eye, resulting in focus at infinity (cycloplegia). The focus was then brought to the appropriate distance (~50 cm) by application of hard (Poly methyl methacrylate) contact lenses, held in place over the cornea by a viscous film (Celluvisc, a carmellose sodium solution). This procedure ensured that the cornea did not dry, and thus remained transparent throughout the experiment. The curvature of the lens was selected to control focus distance, as assessed using an ophthalmoscope. The eye contralateral to the recording site was opened for visual stimulation. The other eye was kept closed.

2.2 Recording Procedure

2.2.1 Recording Set up

For the initial recording of each tract, the location of receptive fields in the visual space was manually determined (by listening to the amplified electrophysiological signals from the electrode) and marked on a translucent hemisphere (Yu & Rosa, 2010). Once the locations were determined, a CRT monitor (Sony Multiscan G520, 100Hz refresh rate) was set up in the region encompassing the receptive fields, 30-50cm from the animal. The location of receptive field was assessed by listening to neural activity while a bar stimulus was moved or flashed across the screens, and drawn on a plastic sheet overlaying a slave monitor visible to the experimenter, which replicated the display presented to the marmoset. The receptive field was mapped in detail for each neuron, and when necessary the position of the monitor was adjusted. Figure 2.2 shows this set-up.

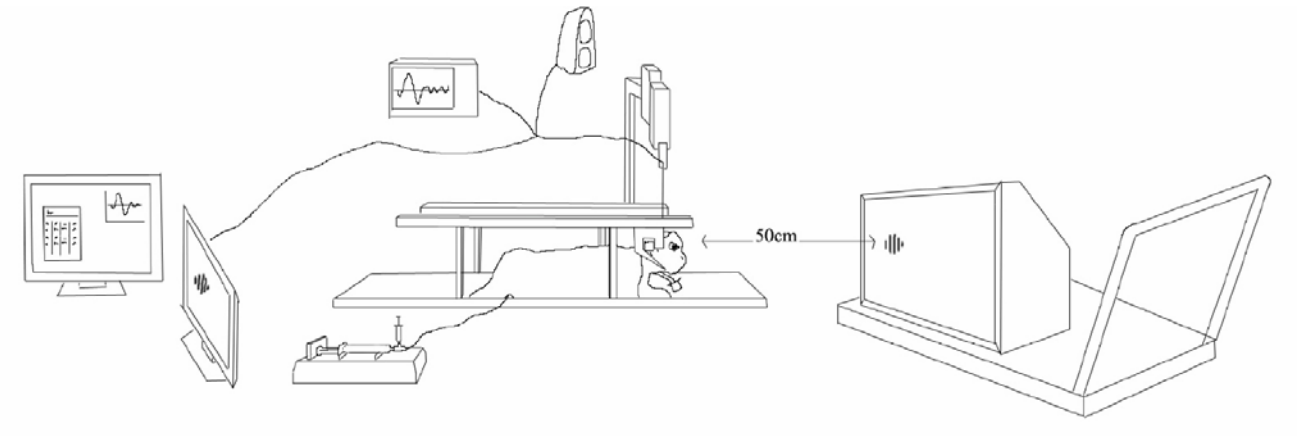

Figure 2.2. Simple schematic of the final set-up for the experiment. A monitor for displaying the stimulus is set up in front of the marmoset, with a mirroring set up in front of the experimenter. Information from the electrode is sent to three separate inputs, two are used for immediate feedback, aiding in isolating receptive fields, a visual and auditory display of the spike waves, and the final input goes to the computer for recording. The screen in front of the marmoset is on a moveable lift to allow for easy adjustment. Anaesthesia is maintained through an intravenous catheter set on automatic pump, as well as a tracheal tube through which nitrous oxide was administered.

Parylene-coated tungsten microelectrodes (~1MΩ), with exposed tips of 10μm were directed towards areas V1 or MT based on stereotaxic coordinates (Paxinos et al., 2012). The electrophysiological signals were amplified and filtered with an amplifier (Model 1800 AC amplifier, A-M Systems), and then sent to a computer (Mac Pro, Quad-Core Intel Xeon 2.8GHz) running the program Expo (designed by Peter Lennie and others), which provided the environment for stimulus generation and data acquisition. Spiking events were sorted on-line using multi-threshold templates (a function provided by Expo).

2.2.2 Visual Stimuli

Within each electrode penetration, neurons were sampled at 100–300 μm intervals. At each location, the boundary of the multiunit minimum response receptive field was carefully mapped using a manually operated stimulus (high-contrast moving bars, small grating patches and flashing light spots) generated on the computer monitor, while the response was monitored through a loudspeaker. Following delineation of the receptive field boundaries, the size of the stimulus was set at 1.2 times the size of the estimated receptive field. The activity

of the largest unit (or units) responding reliably to visual stimulation was isolated while a drifting grating was displayed on the screen, and if necessary, estimates of receptive field boundaries were adjusted. The direction tuning curve of the isolated unit was obtained using a drifting sinusoidal grating, centred on the receptive field, whose SF and TF were adjusted in order to elicit strong responses. For this test, the grating drifted for 2s in one of 12 directions (30^{o} steps), which were presented in a random sequence within each block. After the optimal direction of motion of a neuron was determined, we conducted three series of tests, corresponding to the three stimulus modes described below. These tests were designed to quantify the interaction between responses to different speeds, using stimuli of different SF. In all three tests, high contrast (80%) Gabor patterns with independently varying SF and speed were presented in a 5 × 5 matrix of conditions. The ranges of both parameters were adjusted for each neuron, based on the results of preliminary tests, to ensure that the dynamic ranges were adequately sampled. An additional "blank" condition was included, where a grey screen of average luminance between the high and low luminance peaks of the Gabor pattern was displayed for 2s, to measure spontaneous activity. Each condition was repeated 8 times.

2.2.3 Presentation Modes of the Stimulus

Speed-tuning for each individual neuron was tested with three types of motion stimulus:

The **Standard** mode stimulus is similar to the drifting sinusoidal gratings commonly used in visual research (Baker, 1990; Cavanaugh et al., 2002; Priebe et al., 2003; Yu et al., 2010; Miura et al., 2014), except that drifting Gabor patterns were used instead of the more commonly used drifting grating patches. Drifting Gabor patterns (Fig. 2.3; see also Section 5.2.1. for equations) are created by multiplying a 2D Fourier component (the "carrier") by a 2D Gaussian function (the "envelope"). In contrast to grating patches, whose hard boundaries introduce high frequency components in the Fourier space, the soft-edges of the Gabor

patterns result in a more concentrated distribution of Fourier energy and are more appropriate for this study.

In the **Standard** mode, the Gaussian envelope remained fixed within the neuronal receptive field, while the sinusoidal carrier modulated its phase in time, resulting in localised motion in the optimal direction. Figure 2.3a illustrates this stimulus. However, before the Gabor pattern started to drift, there was a 1s period where the Gabor pattern remained stationary. This was introduced to dissociate the responses of the neuron elicited by the onset of the stimulus from the responses elicited by the motion of the stimulus. The calculations of the response rates (Section 2.3) were based on responses following the time point where the Gabor pattern started to drift.

In the **Moving** mode, the Gabor pattern moved across the receptive field in a straight trajectory in the visual field (Fig. 2.3b). The trajectory was set at 2.5 times the width of the receptive field, allowing for the stimulus to both begin and end outside of the receptive field. The direction of the motion was aligned with the preferred direction of the neuron. As the envelope of the Gabor pattern moved, the carrier remained static relative to the envelope. As in the Standard mode, the stimulus appeared, remaining static for a 1s duration.

An additional difference between the above stimuli was differences in stimulus duration; the Standard mode was presented for a constant two seconds, while the Moving mode was presented for a varying amount of time, depending on the length of the path, and the speed of the stimulus. A third stimulus, called the **Waxing/Waning** mode, was used to examine effects that might be due to the differences in temporal dynamics of the two modes. Similar to the Standard mode, in the Waxing/Waning mode (Fig. 2.3c) the stimulus envelope remained stationary relative to the receptive field, while the carrier wave drifted at a constant speed. However the global contrast was modulated in such a way that the time course of the

modulation of the stimulus at the centre of the receptive field was match to that of a corresponding condition in the Moving mode.

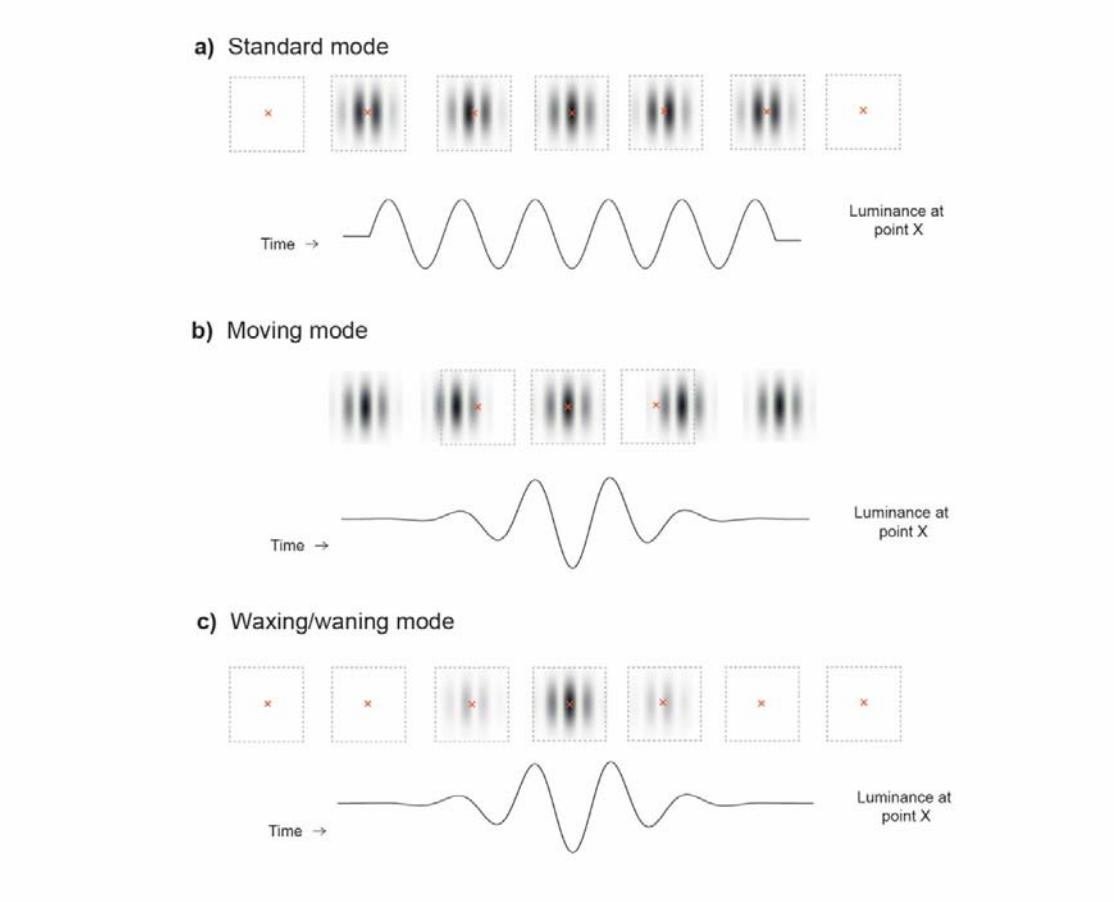

Figure 2.3. The three modes of stimulus presentation. For each panel, the top row illustrates the spatial relationship between the Gabor patch and the receptive field (indicated by the dotted squares) at several time points. The bottom row represents the evolution of stimulus luminance at the centre of the receptive field (marked by a red cross). There was a one second stationary period for Gabor patches in the Standard and the Moving mode, which was not shown in this figure for simplicity.

2.2.4 Fourier Representation of Presentation Modes

Figure 2.4a schematically illustrates how the Fourier space was sampled by the three stimulus modes. In the **Standard** mode, the Fourier energy for each combination of spatial and temporal frequency corresponds to a bivariate Gaussian distribution with major and minor axes parallel to the spatial frequency (SF) and temporal frequency (TF) axes. The SF

bandwidth (determined by the size of the Gabor patch) and the TF bandwidth (determined by the time during which the carrier wave drifted) were constant in all conditions, but as Figure 2.4 is plotted in logarithmic scale, the Gaussian distributions appear to change their width and height as SF and TF increase.

In the **Moving** mode, the Gaussian distributions are tilted 45º in logarithmic scale, parallel to the iso-speed lines (dashed lines). The TF bandwidths varied with speed because the time interval during which the Gabor patch remained within the receptive field varied, but they appear constant in logarithmic scale.

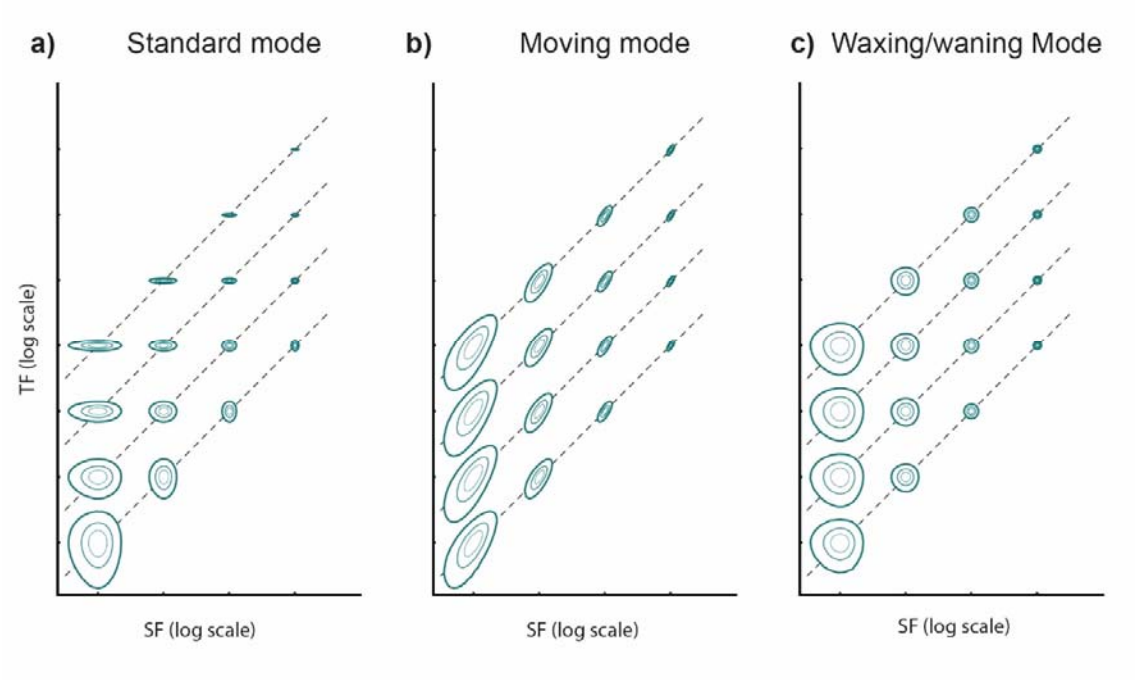

Figure 2.4. Schematic representations of the three stimulus modes in the Fourier space. The energy distributions of the Gabor patches in the three modes are plotted in log-log scale. In this representation, SF and TF combinations that correspond to the same speeds are tilted at 45º angles (iso-speed lines, indicated by dashed lines in the figure). Iso-speed lines closer to the x-axis represent lower speeds. In a) and c) the contours represent the energy distribution of Gabor patterns that drift within a fixed envelope. While b), which is oriented along the iso-speed lines, is the energy distribution of a Gabor pattern which moves across the visual field.

Finally, the energy of a Gabor path in the **Waxing/Waning** mode represents a transition between the two previous conditions. It closely matched that of the Moving mode, as the TF bandwidth varied with speed, due to the variation of the duration of the stimulus. However it was not tilted at a 45° angle, as in the Moving mode, instead showing the same orientation as that in the Standard mode.

2.3 Data Processing and Analysis

2.3.1 Data Analysis

To determine if a neuron is speed-tuned, we used the following procedure:

We estimated the mean stimulus-evoked spike rate for each condition based on the number of action potentials that occurred within a defined time window, adjusted for each neuron. The method for determining this window was designed so that it was time-matched to the onset of response. For each stimulus speed (containing five different spatial frequency conditions) we calculated a space-equalized peri-stimulus time histograms (PSTHs; see Fig. 2.5). This was done by first converting spike times to a spatial scale, which represented the distance travelled by a fixed point in the stimulus (according to $s = v * t$, where s is the spatial distance travelled, t is the time spike and v is the stimulus speed). The space-equalized PSTHs were then fitted with log-transformed skewed-Gaussian functions

$$M(s) = \exp\left[\frac{-(s-s_0)^2}{2(\sigma-\zeta(s-s_0))^2}\right] - \exp\left[\frac{-1}{\zeta^2}\right] \quad (2.1)$$

where s_0 is the point of peak response, σ represents the width and ζ the skewness of the fitted curve. For all modes, the window for calculating spike rates was set to $(s_0 - 1.5\bar{\sigma}, s_0 + 1.5\bar{\sigma})$, where $\bar{\sigma}$ is the mean value of the σ of the fitted histograms for different speeds.

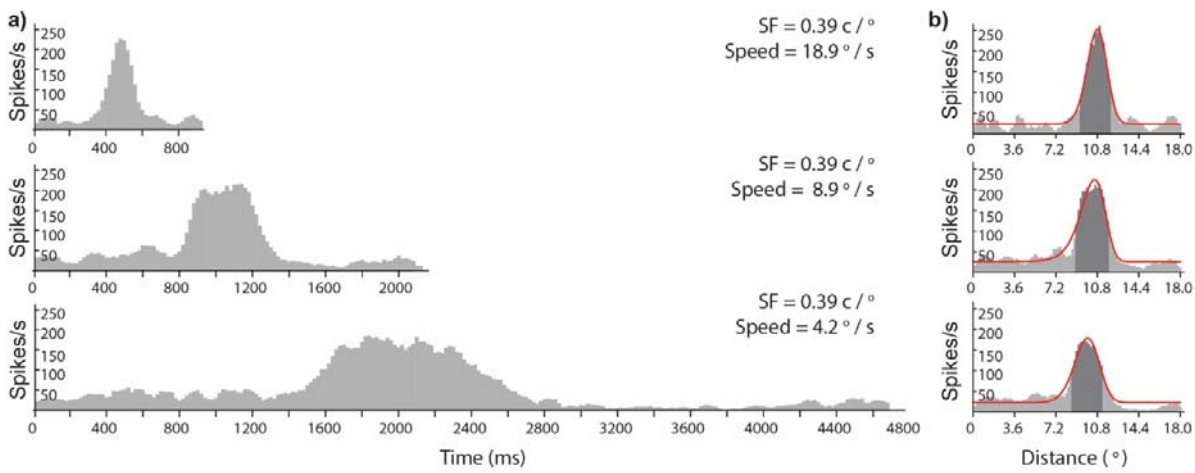

Figure 2.5. The calculation of the space-equalised PSTH for the Moving mode. In a), the PSTHs of a Gabor pattern (central SF = 0.39c/°) moving at three different speeds are plotted. The PSTHs were converted to space-equalised PSTHs in b). The sizes of the bins were adjusted such that the distances travelled by the Gabor pattern were the same length in each bin. The red curves indicate the skewed Gaussian function fitted to the space-equalised PSTHs, and the shaded regions indicate the windows used to calculate the mean response rates. The windows had the same widths (number of bins) but the positions were allowed to vary.

For the Standard mode, all windows were located at the same time-point, which was determined by taking the average of s_0 across all speeds. For the Moving and Waxing/Waning modes, the window was shifted to capture the time intervals where the Gabor pattern was present in the receptive field. This was accomplished by allowing s_0 to vary for different speeds.

The mean spike rates, calculated within the window, produced a tuning surface that represented the neuronal responses as a function of speed and SF.

The SF-speed tuning surface was fitted to two competing models. The first model (M_0) assumed that the neuron has the same optimal speed, regardless of changing spatial frequency, while the second (M_1) assumes that the optimal speed varied systematically with spatial frequency. For M_0, the SF-speed tuning surface was modelled as

$$M_0 = b + c \cdot M_{sf}(sf) \cdot M_{sp}(sp) \tag{2.2}$$

where M_{sf} and M_{sp} were the SF tuning function and the speed tuning function, respectively:

$$M_{sf}(sf) = \exp\left[\frac{-(sf-sf0)^2}{2\left(\sigma_{sf}-\zeta_{sf}(sf-sf0)\right)^2}\right] - \exp\left[\frac{-1}{\zeta_{sf}^2}\right] \quad (2.3)$$

$$M_{sp}(sp) = \exp\left[\frac{-(sp-sp0)^2}{2\left(\sigma_{sp}-\zeta_{sp}(sp-sp0)\right)^2}\right] - \exp\left[\frac{-1}{\zeta_{sp}^2}\right] \quad (2.4)$$

sf and *sp* were spatial frequency and speed transformed by $\log_2$, such that 2^{sf_0} and 2^{sp_0} was the optimal spatial frequency and speed respectively; σ_{sf} and σ_{sp} represent the width of the tuning curves, while ζ_{sf} and ζ_{sp} represent skewness factors ($-2 < \zeta < 2$; zero is perfectly symmetrical). This model had eight parameters.

The second model required the addition of a ninth parameter, Q_{sp}, which quantified how much speed tuning will change with spatial frequency. This is analogous to the Q parameter used in previous studies, for quantifying the change in optimal TF with varying SF (Priebe et al., 2003; Lui et al. 2007a; Miura et al., 2014). The model is expressed as

$$M_1 = b + c \cdot M_{sf}(sf) \cdot M_{sp}(sf, sp) \quad (2.5)$$

where

$$M_{sp}(sf, sp) = \exp\left[\frac{-\left(sp-sp_{sf}\right)^2}{2\left(\sigma_{sp}-\zeta_{sp}(sp-sp0)\right)^2}\right] - \exp\left[\frac{-1}{\zeta_{sp}^2}\right] \quad (2.6)$$

and

$$sp_{sf} = Q_{sp} * (sf - sf0) + tf0 \quad (2.7)$$

In the present convention, if $Q_{sp} = 0$, the optimal speed is independent of SF; in such a case M_1 is equivalent to the SF-speed separable model, M_0 (Fig. 2.6*c*). If $Q_{sp} = -1$, the optimal speed decreases with SF in a pattern that is consistent with perfect TF-tuning (Fig. 2.6a), when $-1 < Q_{sp} < 0$ (Fig. 2.6b), it indicates that the optimal speed decreases with SF, but at a rate lower than expected from a TF-tuned cell.

The Bayesian Information Criterion (BIC) was used to select the model that provided better explanation of the data, taking into account both the complexity of a model and the quality of fit. Formally, BIC is in inverse proportion to the posterior probability of the observed data given a model, and is expressed as $N\left(\frac{RSS}{n}\right) + k * \ln(n)$, where n is the number of data points, k is the number of free parameters in the model and RSS is the residual sum of squares from the estimated model. The probability that M_0 was the better model was given by the evidence ratio:

$$P(M_0|M_0, M_1) = \frac{e^{-BIC0/2}}{e^{-BIC0/2} + e^{-BIC1/2}} \quad (2.8)$$

where BIC_0 was the BIC calculated for M_0, and BIC_1 was the value calculated for M_1. Neurons which had a probability of above 0.5 were considered speed-tuned. This categorisation was also confirmed by a Q_{sp} value of above -0.25, which indicates a low dependence of speed on spatial frequency.

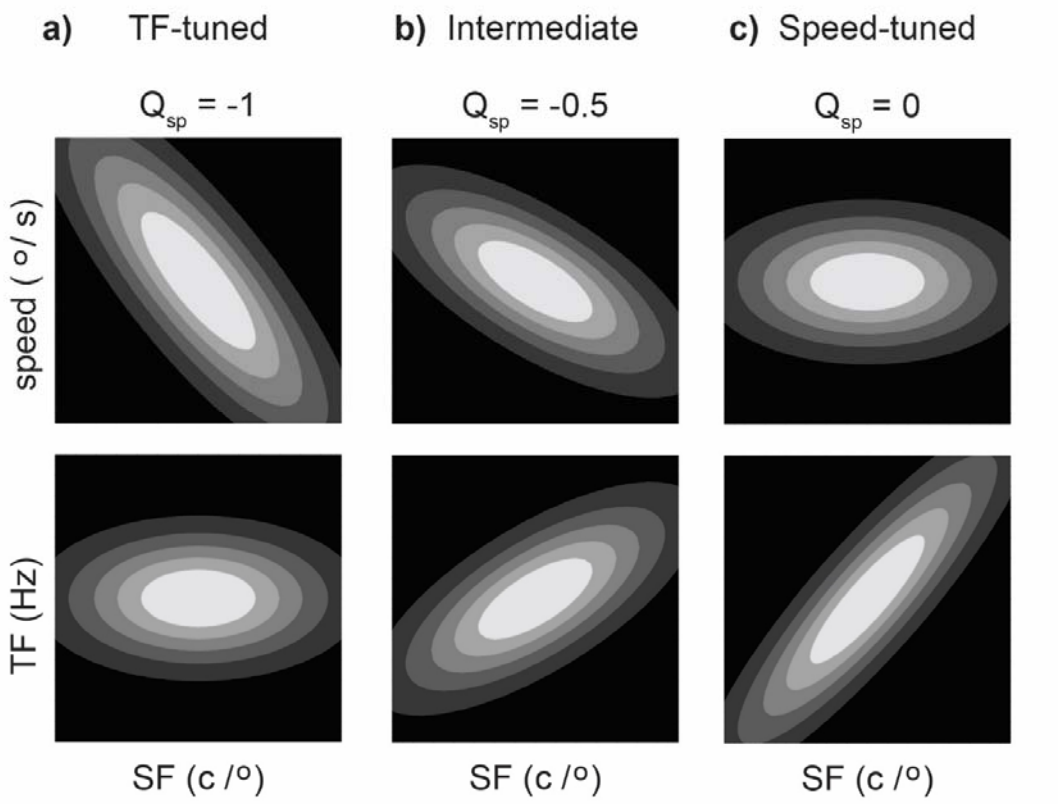

Figure 2.6. Schematic representations of the influence of parameter Q_{sp} in tuning surfaces fitted with model M_1. The Q_{sp} parameter in the SF-speed inseparable model M_1 allows the optimal speed to vary systematically with the SF of the stimulus. In this figure, the contours of the M_1 model at three different levels of Q_{sp} are illustrated. All contours are plotted in the log-log scale. In the upper row, contours are plotted in the SF-speed coordinates used in my experiments, while the lower row shows the same contours plotted in the more commonly used SF-TF coordinates. a) for Q_{sp}= -1, the optimal speed decreased with SF in a pattern that is consistent with perfect TF-tuning. b) for -1<Q_{sp}<0, the optimal speed decreases with SF, but at a rate lower than that expected from a TF–tuned neurons. c) for Q_{sp}=0, the optimal speed is independent of SF. This reflects a perfectly speed-tuned neuron.

2.3.2 Histological Procedures

The electrode tracks were made more readily visible with the aid of small electrolytic lesions (4 μA, 15s), which were placed at various sites during the experiment. At the end of the experiments, the animals were administered an overdose of sodium pentobarbitone and perfused transcardially with 0.9% saline, followed by 4% paraformaldehyde in 0.1M phosphate buffer (pH 7.4). After cryoprotection by increasing concentrations of sucrose, and

sectioning, alternate slides (40 μm) were stained for Nissl substance with cresyl violet, and for myelin (Gallyas, 1979)), allowing reconstruction of electrode tracks and electrolytic lesions relative to histological borders. Area MT was identified by its heavy myelination (Rosa & Elston, 1998). In one animal an alternate series of sections was stained for cytochrome oxidase (Wong-Riley, 1979), which also reveals the boundaries of MT (Yu et al., 2013). Electrode tracks in V1 were identified using well-established criteria, such as the presence of the stria of Gennari.

3. Context-dependent robust coding of stimulus speed in primate extrastriate area MT

As the primary area for the processing of motion, the middle temporal area (area MT) is a likely candidate as a cortical area where the speed of a moving stimulus is unambiguously represented. However, when the hypothesis was examined with drifting sinusoidal gratings, it was reported that only a small population (10% in marmosets, 25% in macaques) of MT neurons are speed-tuned. The majority of MT neurons are instead shown to be tuned to the temporal frequency of the stimulus. In this chapter, I re-evaluated this claim with a new experimental paradigm (as described in the Chapter 2). Parts of the results described in this chapter have been published in Davies et al. (2016).

3.1 Abstract

The ability to estimate the speed of an object irrespective of size, shape or texture is a crucial function of the visual system. However, neurons that show unambiguous (stimulus-invariant) tuning to speed appear to be rare in the Middle Temporal Area (MT), the main motion analysis centre of the primate cortex. Instead, it has been reported that the optimal speed for eliciting responses from the vast majority of MT neurons changes, depending on the spatial frequency (SF) of the stimulus. Here we show that the ability of MT neurons to encode speed has been greatly underestimated, with an ecologically more realistic moving stimulus. Responses of MT neurons were compared in two situations: one in which a sinusoidal pattern moved across the receptive field in a predictable trajectory, versus another in which it was stationed over the receptive field, with motion being defined by a drifting carrier wave (a traditional stimulus in visual physiology). We found that approximately half the neurons showed SF-invariant speed tuning when tested with moving gratings, but not with drifting gratings. These results support the notion that neuronal activity in MT provides accurate coding of speed in natural situations, and that nonlinear operations enable the history and context of visual stimuli to alter how motion information is processed in the brain.

3.2 Introduction

Neurons in the middle temporal area (MT) are regarded as key components of the brain network that processes visual motion (Born & Bradley, 2005), but previous studies have questioned the extent to which their activity can yield reliable information about speed. Using a classical visual physiology paradigm, these studies have used local motion patterns (drifting sinusoidal grating patches), positioned over the receptive fields of MT neurons, to characterize their tuning to speed. Because speed can be derived from the ratio between the temporal frequency (TF) and spatial frequency (SF) of a grating, and given that these two variables can be manipulated independently, this paradigm can be used to determine whether or not a neuron is speed-tuned. Specifically, for a truly speed-tuned neuron, the optimal stimulus TF of a grating is expected to increase with its SF, where the ratio between these variables remains constant (Simoncelli & Heeger, 2001). However, with the exception of one study (Perrone & Thiele, 2001), the results of these experiments have indicated that most MT neurons, like those in earlier stations of the visual system, are tuned to TF, not speed, instead responding optimally to different speeds as spatial frequency changes (Priebe et al., 2003, Lui et al., 2007a, Miura et al., 2014).

There are, however, reasons to question whether earlier studies have provided a full picture of the ability of MT neurons to signal speed. The motion signals in drifting gratings originate from the temporal modulation of the spatial phases of localized carrier waves, but the stimulus envelope remains stationary over the same retinal locations. With rare exceptions (Laan et al., 2014), this combination of features is not found in natural images, and therefore it can be argued that the primate brain may not have evolved to optimally evaluate speed in such situations.

That the behaviour of motion processing neurons can be characterized with drifting gratings has been justified, on theoretical grounds, by the fact that the images of objects that travel in

straight trajectories are represented in Fourier space as oriented lines or planes, and can therefore be captured by linear filters (Watson & Ahumada, 1985; Simoncelli & Heeger, 1998). This being the case, the responses of neurons that process motion using linear summation of signals could be adequately characterized using drifting gratings. However, the summation linearity assumption has already been challenged by the demonstration that MT neurons show greater evidence of speed-tuning when stimulated with two superimposed gratings of different SFs (Priebe et al., 2003), a result that established that the Fourier components of a moving pattern can interact nonlinearly to alter information processing in the primate brain.

Translation of an object across visual space affects the exact combination of spatial and temporal frequencies to which the visual system is exposed, thus also creating opportunity for nonlinear interactions in neural processing. However, it has not yet been tested whether this type of moving pattern elicits a greater degree of SF-invariant speed-tuning in visual cortex. To test if this is the case, in the present study we compared the responses of MT neurons to drifting Gabor patches (sinusoidal gratings weighted by a Gaussian envelope) and Gabor patches that moved across the receptive field, passing consecutive retinal locations. The use of the latter type of stimulus allowed us to incorporate motion trajectory, an important feature of natural motion signals, while retaining the ability to vary SF and TF in a precise manner. With this method, we show that stimuli that move across the receptive field of a neuron are processed differently from locally drifting stimuli, and that only the first elicit robust speed-tuning in a large number of MT cells.

3.3 Methods

Single unit recordings were obtained from area MT in 3 adult New World monkeys (*Callithrix Jacchus,* the common marmoset). Experiments were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes, and all procedures were approved by the Monash University Animal Ethics Experimentation Committee.

Preparation of the animal for *in vivo* recordings of neuronal responses involves the use of several surgical stages, based on the protocol described in Bourne and Rosa (2003). This procedure is described in full detail in the Methods chapter.

Within each electrode penetration, neurons were sampled at 100-300 μm intervals. At each location, the boundary of the multiunit minimum response receptive field was carefully mapped using a manually operated stimulus (high-contrast moving bars, small grating patches and flashing light spots) generated on the computer monitor, while the response was monitored through a loudspeaker. Following delineation of the receptive field boundaries the activity of the largest unit (or units) responding reliably to visual stimulation was isolated while a drifting grating was displayed on the screen and, if necessary, estimates of receptive field boundaries were adjusted.

The direction tuning curve of the isolated unit was obtained using a drifting sinusoidal grating centred on the receptive field, whose SF and TF were adjusted in order to elicit strong responses. For this test the grating drifted for two seconds in one of 12 directions (30^o steps), which were presented in a random sequence within each block.

3.3.1 Stimulus

After the optimal direction of motion of a neuron was determined, we conducted three series of tests, corresponding to the three stimulus modes described in the Methods section (2.2.2.1).

The **Standard** mode replicated the typical way in which drifting gratings were presented in earlier experiments that examined speed-tuning in visual cortex: the sinusoidal carrier of a Gabor patch drifted within a stationary envelope centred on the cell's receptive field, for a fixed time interval (Fig. 2.3a). However, the use of a Gabor patch instead of a traditional grating with sharp boundaries eliminates the possible confound represented by richer combinations of spatial frequencies at the edges of the stimulus.

In the **Moving** mode, a non-drifting Gabor patch appeared well outside the neuron's receptive field, and then moved in a straight trajectory, passing across the receptive field (Fig. 2.3b).

The **Waxing/Waning** mode was designed to mimic the effect of viewing a moving stimulus through an aperture corresponding to the receptive field: the Gabor patch drifted within a stationary envelope, as in the Standard mode, but the global contrast of the Gabor pattern was temporally modulated in such a way that the time course of the modulation at the centre of the receptive field was matched to that of the corresponding stimulus in the Moving mode (Fig. 2.3c).

In all three modes, the sinusoidal components of the Gabor pattern moved at a constant speed at each point in the cell's receptive field, but only in the Moving mode did the envelope of the pattern move across the visual field. In all tests, the diameter of the Gabor patch was set to be 10% larger than that of the receptive field (carefully mapped with moving bars and small gratings), to further reduce the possibility of boundary effects, and the direction of motion was set to the neuron's optimal, determined quantitatively.

3.4 Results

Given the assumption that the integration of motion signals by MT neurons primarily involves a linear summation stage (Watson & Ahumada, 1985; Simoncelli & Heeger, 1998; Rust et al., 2006; Bradley & Goyal, 2008; Nishimoto & Gallant, 2011), the three stimulus presentation modes would not be expected to have significant effects on the degree of speed tuning (See Chapter 5 for a formal argument). To test this prediction, we characterized the response properties of 110 neurons in three adult marmosets, confirmed to be in area MT by histological reconstruction (Rosa & Elston, 1998). The receptive fields were in the near-peripheral visual field, between 10˚ and 30º eccentricity, and encompassed both the upper and lower contralateral quadrants. The diameters of the receptive fields were in the range of 5º to 15º (median value 10.2º).

3.4.1 Single-neuron response patterns

Responses of most neurons fell into two main categories: neurons that did not change speed-tuning classification across any condition, and neurons that were speed tuned only when presented with the Moving mode stimulus. Figure 3.1 illustrates data obtained in a representative neuron of the latter category; whereas Figure 3.2 illustrates data obtained in a representative neuron of the first category.

In Figures 3.1 and Figure 3.2, the firing patterns of the neuron when stimulated by Gabor patterns of 25 spatial frequency/speed combinations in the Standard mode (Fig. 3.1a, Fig 3.2a) and in the Moving mode (Fig. 3.1b, Fig 3.2b) are shown as space-equalized peri-stimulus time histograms (Section 2.3.1). The window used for calculating the firing rates for each condition is indicated by the vertical black lines.

To facilitate visualisation, the “tuning surfaces” are plotted as rasters, where the average firing rates are represented in grey scale (high firing rate is displayed as a lighter colour).

These tuning surfaces are what were used for determining if a neuron was speed-tuned (using the procedure described in Section 2.3.1), but speed-tuning can be most easily comprehended by plotting the speed-tuning curves for Gabor patterns of individual spatial frequencies (Fig. 3.1 & Fig. 3.2, bottom row). For the non-speed-tuned responses, the peak of speed-tuning curve varies according to spatial frequency, (Fig. 3.1e, Fig. 3.2e, Fig. 3.2f), whereas for speed-tuned responses, the neuron's preferred speed remains stable despite changing spatial frequency (Fig. 3.1f).

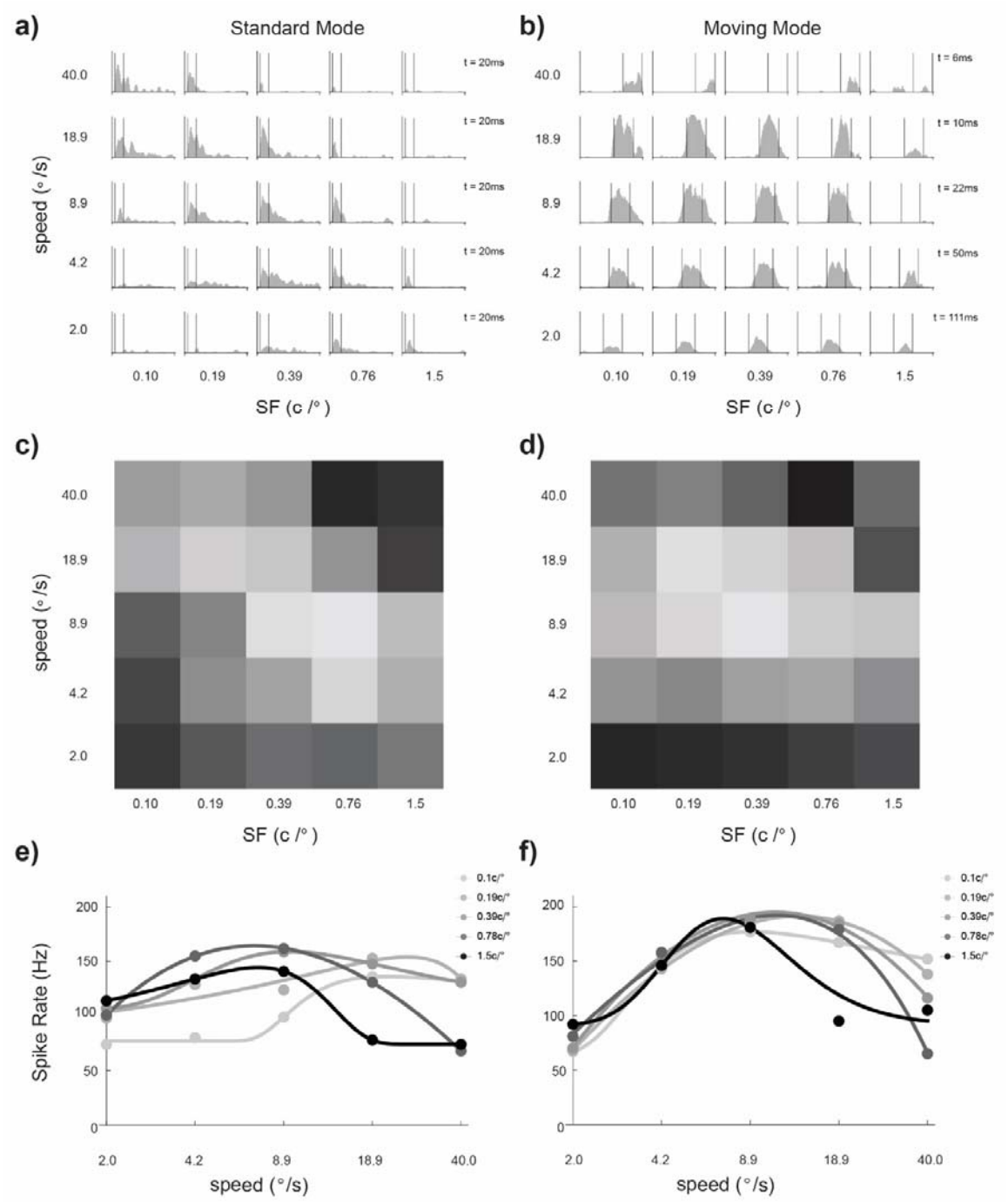

Figure 3.1. Sample neuronal responses from a single neuron that was speed-tuned in the Moving mode. Columns show responses to the Standard (a, c, e) and Moving (b, d, f) modes. For each column, the first row shows the responses to each SF-speed combination as a PSTH, adjusted for equal representation of space (the width of the bins for each speed are indicated on the right of each PSTH row). The window of time used for analysis is indicated by the two vertical lines. The second row is a visual representation of tuning. Shade correlates to the rate of firing, such that a lighter shade represents a higher rate of firing. The third row displays speed-tuning curves obtained using Gabor patches of different SF. The circles represent the measured response rates, and the curves are fitted skewed log-Gaussian functions.

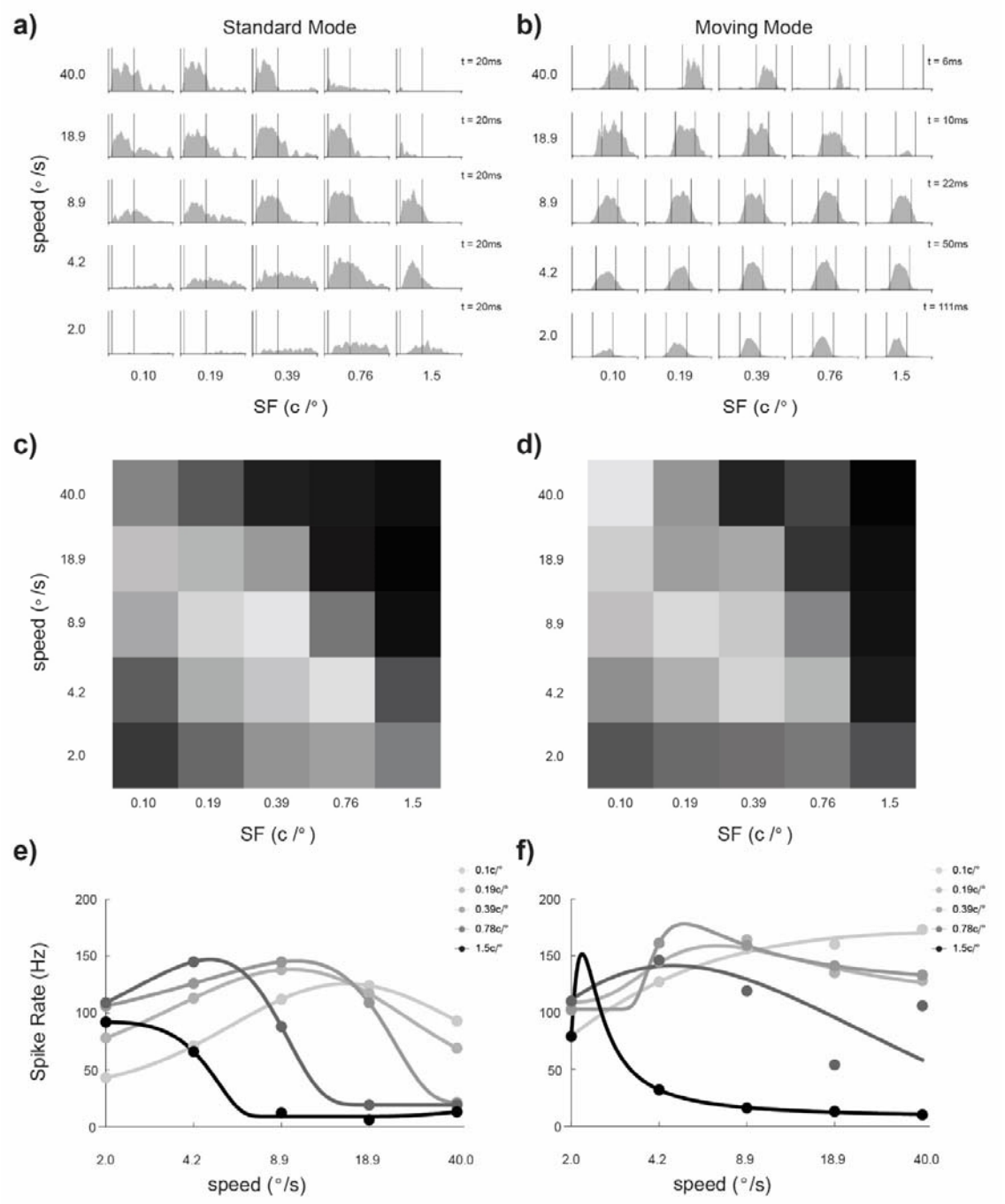

Figure 3.2. Sample neuronal responses from a single neuron that was speed-tuned in neither condition. The data are presented in the same format as Figure 3.1.

Using the procedure detailed in Section 2.3.1, the tuning surfaces were then fitted with two models: an SF-speed separable model (M_0, Fig. 3.3, middle row), which assumed that the SF and speed-tuning curves were independent of each other, and an SF-speed inseparable model (M_1, Fig. 3.3, bottom row), in which the optimal speed could vary systematically with the stimulus SF. We used the Bayesian Information Criterion (BIC, Yu et al., 2010) to calculate $P(M_0|M_0,M_1)$, the probability that M_0 provided a better fit to the data over M_1.

For the neuron illustrated in Figure 3.1, $P(M_0|M_0,M_1)$ was lower than 0.001 for stimulation in the Standard mode (Fig 3.3, left col), indicating that there was more likely a systematic relationship between the optimal speed and SF of the stimulus (M_1). However, when the same neuron was tested with the Moving mode stimulus (Fig 3.3, right col.), $P(M_0|M_0,M_1)$ was 0.91, indicating that the pattern of responses was best described by a model according to which the neuron's optimal speed was not influenced by the spatial frequency of the stimulus (M_0).

The Q_{sp} parameter obtained in model M_1 indicates the dependency between spatial frequency and speed (Section 2.3.1, Fig. 2.6). The estimated Q_{sp} for the neuron shown in Figure 3.3 were -0.73 and -0.09 for the Standard and Moving mode respectively, where the higher Q_{sp} value for the Moving mode is an additional indicator that the neuron was not influenced by spatial frequency. In contrast, for the neuron illustrated in Figure 3.4, the estimated Q_{sp} values were -0.90 and -0.71 for the Standard and Moving modes respectively.

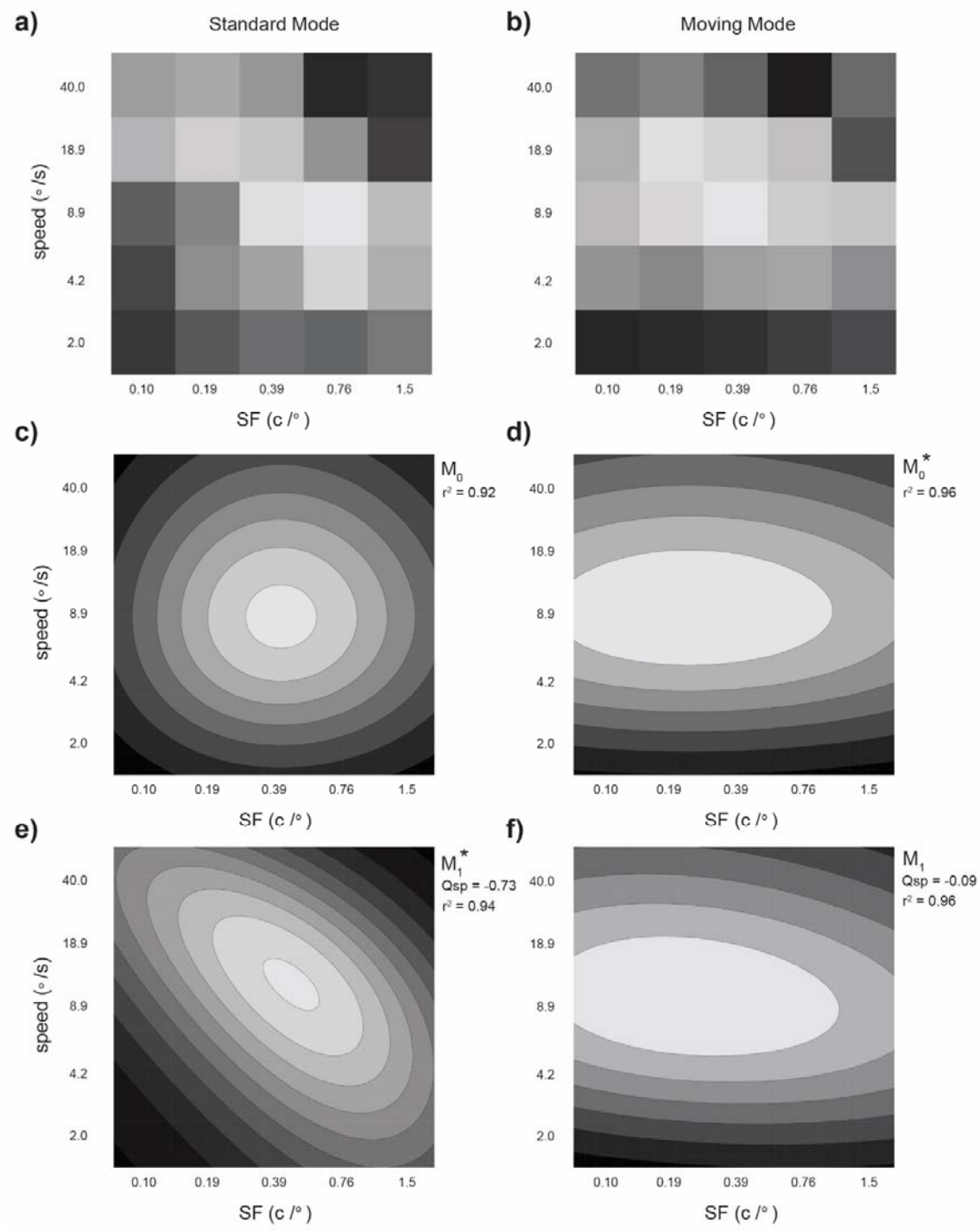

Figure 3.3. Example of neuronal response properties from one neuron fit to M_0, the speed tuned model, and M_1, the non-speed tuned model, for a neuron that was speed-tuned in the Moving mode. The results from both the Standard (a, c, e) and the Moving (b, d, f) mode are shown. The top row shows a raster plot that represents conceptually which spatial frequency/speed combination resulted in the highest firing rate of the neuron, with a larger diameter indicating a higher response rate. The middle row shows the data fit to M_0, while the bottom row shows the data fit to M_1. BIC calculations results in a value of $P(M_0|M_0,M_1)$ to determine the probability that M_0 is a better fit. For the Standard mode, $P(M_0|M_0,M_1) < 0.001$ indicating a non-speed tuned neuron. The Moving mode neuron was found to be speed-tuned with a $P(M_0|M_0,M_1)$ of 0.91.

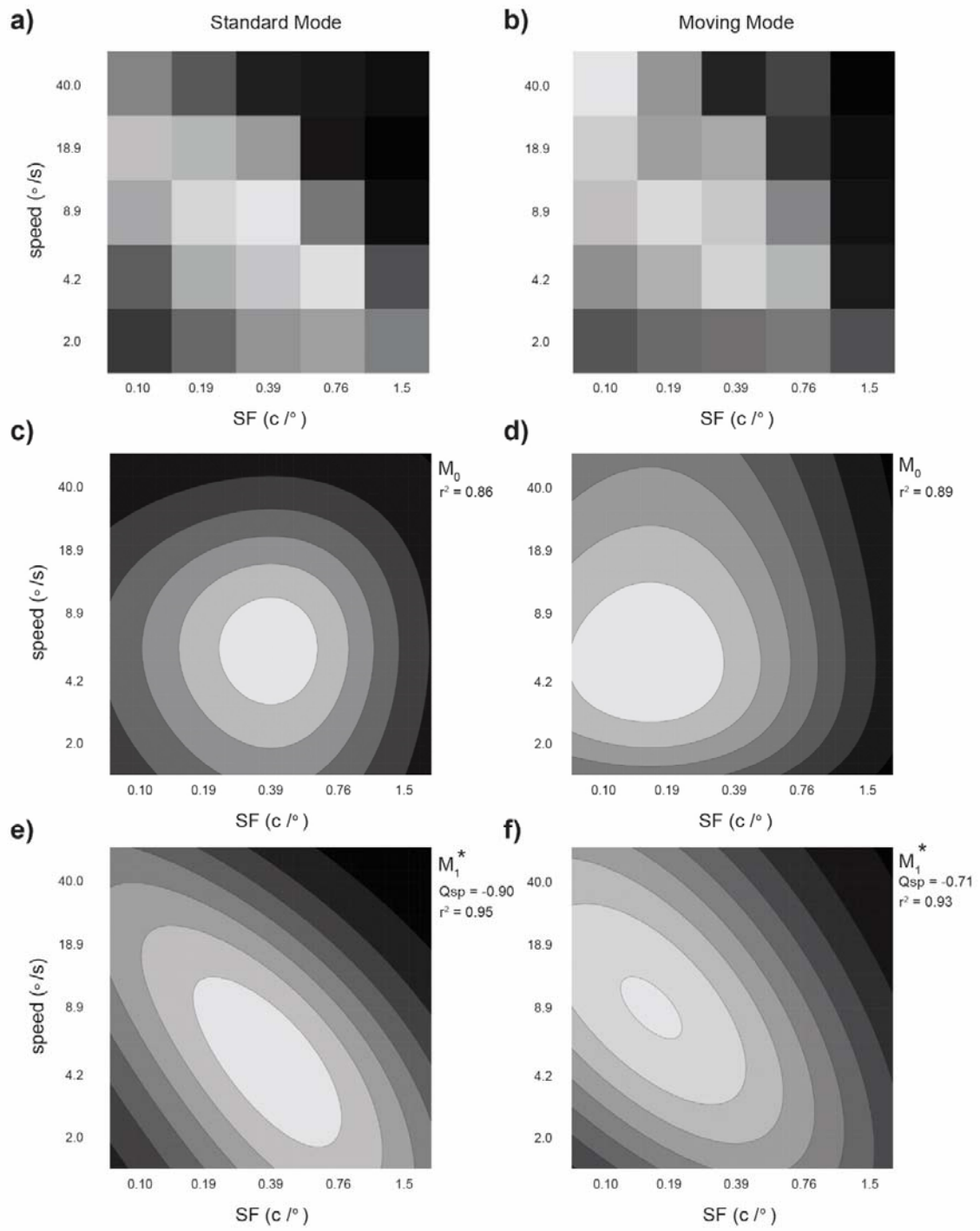

Figure 3.4. Example of neuronal response properties from one neuron fit to M_0, the speed tuned model, and M_1, the non-speed tuned model, for a neuron that was not speed-tuned in either condition. The data are presented in the same format as Figure 3.3.

3.4.2 Speed Tuning Population Statistics

The median $P(M_0|M_0,M_1)$ values for the Standard and Moving modes were significantly different, with a median of 6.07E-6 and 0.74 respectively (Wilcoxon test: V = 495, P<0.0001). These distributions are plotted in Figure 3.5, left column. The distribution for the Standard mode was unimodal and heavily skewed towards 0, in contrast, the distribution for the Moving mode was found to be bimodal (Dip test: D = 0.181, p<0.0001) showing a clear increase in the number of neurons for which M_0 provided better explanation of the data.

In the right column of Figure 3.5, we compare values of Q_{sp} that were obtained from M_1. Consistent with previous studies (Priebe et al., 2003; Lui et al. 2007a; Miura et al., 2014), these distributions were unimodal, and compatible with the notion that there is a continuum between speed-tuning and TF-tuning among MT neurons. Crucially, the distributions were shifted according to the mode of presentation of the stimulus. Median values were -0.92 and -0.33 for the Standard and Moving mode, respectively, indicating that the average dependency between SF and speed across the sample was lower in the Moving mode. The populations were significantly different from each other (Wilcoxon test: V = 566, P<0.0001).

Using a threshold criterion of $P(M_0|M_0,M_1)>0.5$, we found that only 1 cells (0.9%) was classified as speed-tuned when tested with Gabor patches presented in the Standard mode. However, this estimate increased dramatically to 53 (48.2%) when stimuli in the Moving mode were used (χ^2=66.4, p<0.0001). As an alternative method for classifying speed-tuning, we also counted the number of neurons whose M_1 model yielded Q_{sp} values between -0.25 and 0.25. This criterion was chosen with the heuristics that the fitted M_1 surfaces with Q_{sp} values within this interval were often very similar to the fitted M_0 surface. The numbers of speed-tuned cells according to this criterion were 1 (0.9%) and 51 (46.4%), according to tests conducted in the Standard and Moving modes respectively, thus confirming that our findings were not simply due to the method of classification. These values were, again, significantly

different (Standard vs. Moving: χ^2=62.96, p<0.0001. In summary, only tests using stimuli presented in the Moving mode revealed a large population of speed-tuned cells in MT, irrespective of the classification criterion used.

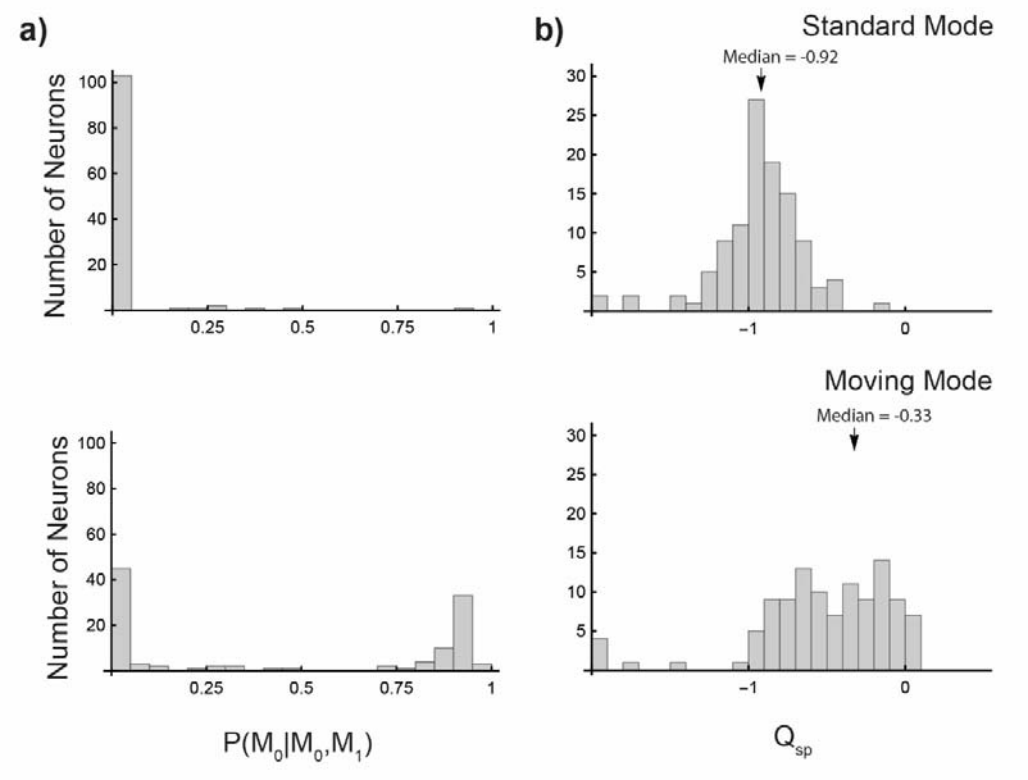

Figure 3.5. Distributions of $P(M_0|M_0,M_1)$ and Q_{sp} under the stimulation of the Standard mode (top row) and the Moving mode (bottom row). a) Histograms showing the distributions of $P(M_0|M_0,M_1)$. b) Histograms showing the distributions of Q_{sp}, the arrows indicate median values.

3.4.3 Effects of Presentation Time

The main limitation inherent in comparing the Standard and the Moving modes was the factor of time. The Standard mode was presented for a constant duration of two seconds, while the Moving mode ran for as long as it took for the Gabor pattern to complete its trajectory. In order to account for this difference, the Waxing/Waning mode was used. This was a drifting Gabor like that of the Standard mode and activated only a small area of the visual field.

However, the Waxing/Waning mode appeared for the same duration as the Moving mode, with a luminance that was modulated over time such that it reflected that which was seen in the centre of the receptive field in the corresponding condition of the Moving mode. The distributions of $P(M_0|M_0,M_1)$ and Q_{sp} are plotted in Figure 3.6.

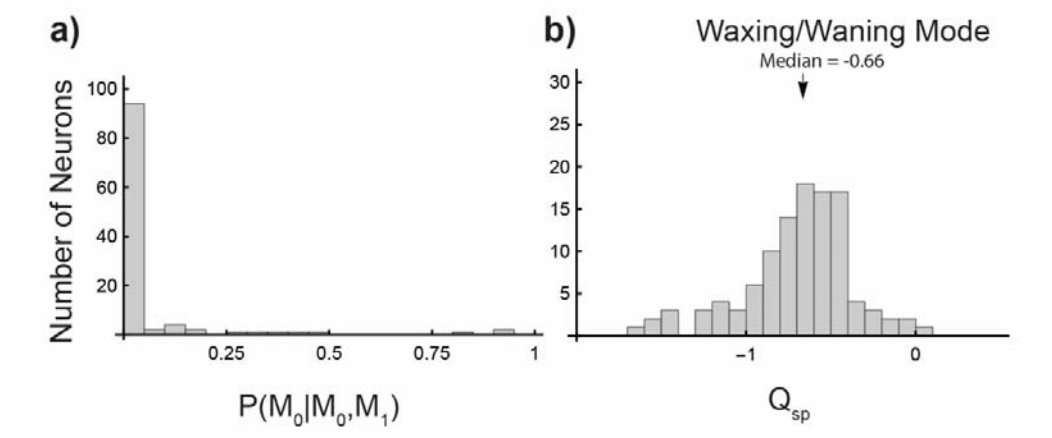

Figure 3.6. Distributions of $P(M_0|M_0,M_1)$ and Q_{sp} under the stimulation of the Waxing/Waning mode. a) Histogram showing the distribution of $P(M_0|M_0,M_1)$. b) Histogram showing the distribution of Q_{sp}, median value is indicated by the arrow.

In this condition, there were 3 (2.7%) speed-tuned neurons. This result was not significantly different from the Standard mode (χ^2=1.02, p=0.3129), but was found to be significantly different from the Moving mode (59.89, p<0.0001). However when examining the $P(M_0|M_0,M_1)$ and Q_{sp} values using Wilcoxon signed-rank tests, the Waxing/Waning mode was significantly different from both of the other modes; though it was more similar to the Standard mode than the Moving mode (Wilcoxon test: Standard mode: $P(M_0|M_0,M_1)$, V = 1978, p = 0.0014; Q_{sp}, V = 1752, p<0.0001; Moving Mode: $P(M_0|M_0,M_1)$, V = 5350, p<0.0001; Q_{sp}, V = 5079, p<0.0001).

3.4.4 Changes of Speed-Tuning Classification Across Stimulus Modes

The population analyses described above prompt the question of whether there was a consistent pattern of change in speed-tuning classification, at the level of individual neurons.

As summarized in Table 3.1, we found that 58 neurons (52.7% of the sample) did not change their speed-tuning classification, irrespective of stimulus mode. Most of these neurons were consistently non-speed-tuned with one neuron that was consistently speed-tuned across all three modes. Among the cells that did change classification, there were no neurons that were classified as speed-tuned in the Standard mode, but became non-speed-tuned in the Moving mode. Rather, many neurons that were non-speed-tuned with Standard mode stimuli became speed-tuned with Moving mode stimuli (52, 47.3% of the sample). Of these, there were 50 neurons which were also not speed-tuned when presented with stimuli in the Waxing/Waning mode.

Table 3.1. Speed-tuning classification of neurons across the three test conditions. Speed-tuning was determined using a threshold criterion of $P(M_0|M_0,M_1) > 0.5$.

Standard Mode	**Moving Mode**	**Waxing/Waning Mode**	**Number**
Speed-Tuned	Speed-Tuned	Speed-Tuned	1 (0.9%)
Speed-Tuned	Non Speed-Tuned	Speed-Tuned	0 (0%)
Speed-Tuned	Speed-Tuned	Non Speed-Tuned	0 (0%)
Speed-Tuned	Non Speed-Tuned	Non Speed-Tuned	0 (0%)
Non Speed-Tuned	Speed-Tuned	Speed-Tuned	2 (1.8%)
Non Speed-Tuned	Non Speed-Tuned	Speed-Tuned	0 (0%)
Non Speed-Tuned	Speed-Tuned	Non Speed-Tuned	50 (45.5%)
Non Speed-Tuned	Non Speed-Tuned	Non Speed-Tuned	57 (51.8%)

We also examined how estimates of the parameter Q_{sp} change for individual neurons, with respect to the stimulus presentation mode (Fig. 3.7, upper row). We found that, for the vast majority of neurons (93, 84.5%), changing the stimulus from the Standard mode to the Moving mode resulted in an increase in Q_{sp} (Fig. 3.7a), and that a similar marked change occurred from the Waxing/Waning to Moving mode (Fig. 3.7a; 75 neurons, 68.2%). Two features about these plots were noteworthy. First, among the neurons that became speed-tuned (indicated by white dots in the scatter plots), the Q_{sp} values obtained in the comparison modes (represented in the abscissas) were widely distributed. The change was therefore not explained by a "tip of the iceberg" effect, where an equal amount of tilt in the SF-speed space was introduced by the moving stimulus, causing nearly speed-tuned neurons in the Standard and Waxing/Waning modes to become speed-tuned. Second, there were no obvious clusters. It is therefore unlikely that the change in Q_{sp} was mediated by a mechanism that operates on specialized populations. To quantify this effect, we plotted the difference between Q_{sp} values obtained with different stimulus modes (Fig. 3.7, lower row). The median values were 0.50, 0.23, and 0.26 in the Standard to Moving, Waxing/Waning to Moving and Standard to Waxing/Waning comparisons, respectively. All distributions were unimodal**.**

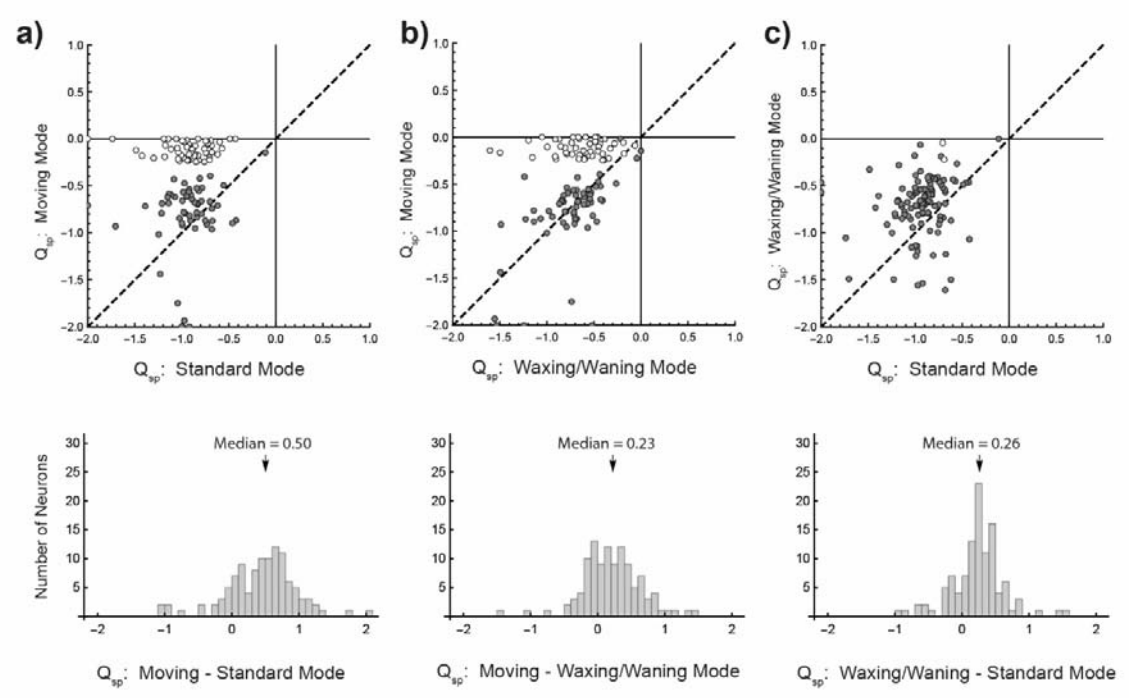

Figure 3.7. Changes in Q_{sp} between two modes. The upper row compares the values of Q_{sp} in tests conducted in the a) Standard and the Moving modes b), the Waxing/Waning and the Moving modes and c), the Standard and the Waxing/Waning modes. White circles represent neurons that were changed from being non-speed-tuned (in the mode specified on the x-axis) to being speed-tuned (in the mode specified on the y-axis). Black circles represent all other neurons. The dashed lines represent values where Q_{sp} was identical in the two modes. The lower row illustrates the distributions of the difference of Q_{sp} in different modes.

3.4.5 Spatiotemporal Tuning Parameters

The scaling parameters of the best-fitted model (b and c in equations 2.2 or 2.5) quantify the range of firing rate of the neuron. As determined by a two-way ANOVA, there was no significant effect of stimulus mode on the interaction or variation of these variables ($F(1,220) = 0.0$, $P>0.9999$), suggesting that the maximum firing rate of the neuron is not affected by the stimulus presented to the receptive field.

3.4.5.1 Spatial Frequency Tuning

The median values of the optimal spatial frequency were 0.31 c/^{o}, 0.24 c/^{o} and 0.25 c/^{o} respectively for the Standard, Moving, and Waxing/Waning modes (Fig. 3.8, left). The differences are small, but pairwise comparison indicated that the optimal spatial frequency was significantly higher when the neuron was stimulated by the Standard mode stimulus when compared to the other two modes (Wilcoxon test: V= 3801, P=0.0257; V = 4531, p<0.0001). No significant difference was found between the Moving and the Waxing/Waning mode (Wilcoxon test: V= 3801, P=0.0257; V = 4531, p<0.0001).

The spatial frequency tuning curve is determined by two other parameters: σ_{sf} characterises the bandwidth of the tuning curve and ζ_{sf} the skewness. For bandwidth (σ_{sf}), there was a significant difference between the Standard (1.59) and Waxing/Waning mode (1.80; Sign Test, Z = 37, p =0.012) where the Waxing/Waning mode was found to have a slightly larger bandwidth. There was no significant difference in ζ_{sf} (Median values = 0.01 and 0 respectively; Sign Test, Z = 31, p = 0.4). Between the Moving and the Waxing/Waning mode, there was no significant difference between either bandwidth (1.64; Sign Test, Z = 37, p = 0.10) or skewness (0; Sign Test, Z = 46, p = 0.08). Interestingly, neither the bandwidth (Sign Test, Z = 52, p = 1) nor the skewness (Sign Test, Z = 47, p = 0.15) was significantly different for the Standard and the Moving mode, indicating that when the stimulus was changed from the Standard to the Moving mode, the spatial frequency tuning curves were shifted towards a smaller value while maintaining their shape.

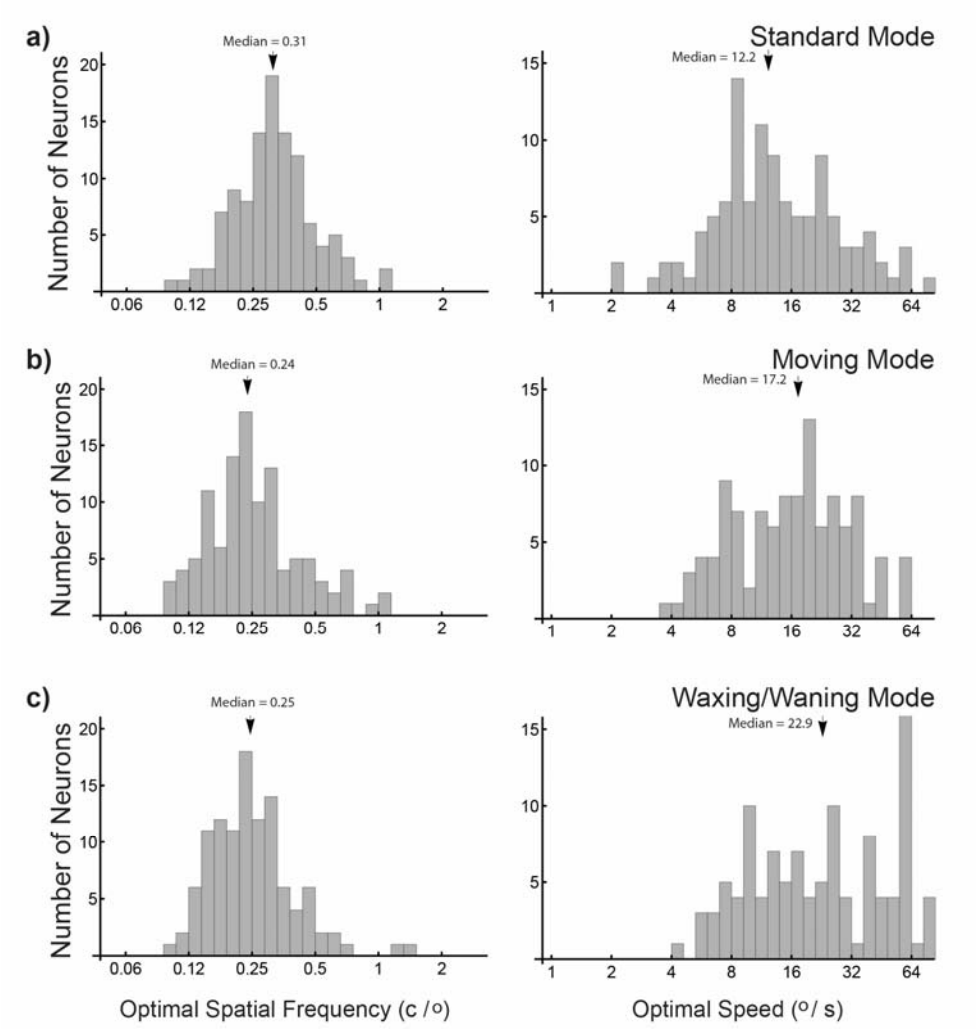

Figure 3.8. Distributions of optimal spatial frequency and speed under the presentation of the three stimulus modes: a) Standard, b) Moving and c) Waxing/Waning modes. Left column: Histograms showing the distributions of optimal spatial frequency, the arrows indicate median values. Right column: Histograms showing the distributions of optimal speed, arrows indicate median values.

3.4.5.2 Speed Tuning

The median optimal speed for the Standard, Moving, and Waxing/Waning modes were 12.2 °/s, 17.2 °/s and 22.9 °/s respectively (Fig. 3.8, right). Pairwise comparison indicated that the optimal speed was significantly lower in the Standard mode than in the Moving (Wilcoxon test, $V = 2086$, $p = 0.0040$) and Waxing/Waning mode (Wilcoxon test, $V = 728$, $p < 0.0001$). The Waxing/Waning mode was found to have the highest preferred speed of all the three stimuli (Wilcoxon test: Moving vs Waxing/Waning mode, $V = 5407$, $p < 0.0001$).

For the bandwidth (σ_{sp}) and skewness (ζ_{sp}) variables of the speed tuning curve, only minor differences were found. Median values for bandwidth were 1.61, 1.5 and 1.6 respectively, and there was no significant difference in bandwidth between any of the conditions (Sign test: Standard vs Moving, Z = 56, p = 0.63; Standard vs Waxing/Waning, Z = 56, p = 0.56; Moving vs Waxing/Waning, Z = 51, p = 0.77). Median values of skewness were 0.02, 0.02 and 0.05 respectively. Despite median values pairwise comparison showed no significant difference in skewness between the Waxing/Waning mode to the other two modes (Sign test: Standard, Z =29, p 0.08; Moving, Z = 23, 0.49), however there was a significant difference between the Standard and Moving mode (Sign Test, Z = 24, p = 0.0013).

3.5 Discussion

Characterization of the responses of MT cells to moving visual patterns is widely regarded as key to our understanding of the brain mechanisms involved in visual motion processing (Born & Bradley, 2005) .We have addressed an important aspect of this question: to what extent can MT neurons signal the speed of a moving object in an unambiguous manner?

Previous single-unit recording studies in both awake and anesthetized animals have reported that only a minority of neurons in area MT (10-25%) yielded responses that reflect tuning to the speed of the carrier wave of a grating, irrespective of its SF (Priebe et al., 2003, Lui et al., 2007a, Miura et al., 2014). Refining on these studies, the present experiments used Gabor patches as visual stimuli, instead of sharply delimited gratings, to avoid the possibility of inadvertent stimulation of the neuron's receptive field by other spatial and temporal frequencies (which could affect estimates of speed-tuning; Priebe et al., 2003; Lui et al., 2007a). In tests where Gabor patches were presented in the Standard (drifting) mode, we found an even lower proportion (~1%) of speed-tuned cells in MT, in comparison with previous work using traditional gratings (including a study that used otherwise identical methods; Lui et al., 2007a). However, our key observation here is that many of these same cells, amounting to approximately half of the tested MT population, revealed SF-invariant speed tuning in tests where the Gabor patch moved along a trajectory that crossed the receptive field. Our results suggest a form of spatiotemporal contextual effect, in which the prior history of the stimulus influences how information is processed in MT.

Many computational models of motion processing in MT (Watson & Ahumada, 1985; Simoncelli & Heeger, 2001; Rust et al., 2006; Bradley & Goyal, 2008; Nishimoto & Gallant, 2011) incorporate nonlinear elements such as static linearity and divisive normalisation, but integrate activities across spatiotemporal channels linearly. Under the linear pooling assumption, a neuron's response profile to the speed of a moving stimulus would be

predictable from its responses to drifting gratings, despite the nonlinear elements in the model. The fact that approximately half of the neurons in our sample changed classification, from non-speed-tuned in the Standard mode to speed-tuned in the Moving mode, indicates significant nonlinear interactions among spatiotemporal channels in a large population of MT neurons (See Chapter 5 for a formal argument). Psychophysical experiments using stimuli consisting of the sum of two gratings have reported nonlinear summation behaviour in speed judgments (Smith & Edgar, 1991). In addition, Priebe et al. (2003) superimposed two drifting gratings with different SFs but the same speed, and reported that responses of neurons in macaque MT were more tuned to speed of the superimposed gratings than the predicted responses patterns as measured by single drifting gratings. Our findings are compatible with Priebe et al. (2003), because the Moving mode stimulus contained a richer set of spatiotemporal frequencies, compatible with the spatial displacement of the envelope of the Gabor pattern in a trajectory. Importantly, we further showed that although nonlinear summation can alter how motion signals are processed in MT, the effect was observed only in about half of the population. Thus, TF-based computation is still an important component in the cortical processing of motion signals.

The novel aspect of the Moving mode stimulus is that the Gabor pattern travelled across the receptive field. An interesting possibility is that the temporal context of the moving trajectory allows the history of the stimulus outside the receptive field to influence how information is processed when the stimulus is in the receptive field. Such mechanisms have been suggested by psychophysical studies (Mckee & Welch, 1985; Grzywacz et al., 1995; Roach et al., 2011; Neri, 2014). Further studies are required to test this possibility.

In addition to the Moving mode, we also used a variant of drifting stimuli (Waxing/Waning mode) in which the duration of presentation of the Gabor patch was consistent with that of a Moving mode stimulus through an aperture. The varying duration (and therefore the varying

TF bandwidth) of the stimulus in this presentation mode was found to have no significant effect on the proportion of speed-tuned neurons in MT, in comparison to the Standard mode. This result suggests that dependence on the temporal properties of a stimulus is not marked, in terms of determining the degree of speed-tuning.

The Waxing/Waning mode was found to prefer significantly higher speeds (median = 22.9°/s) when compared to the Standard and Moving modes. It is possible that this finding might have some significance to information processing in MT. But alternatively, it is also possible that it is an artefact of the stimulus design. At very fast speeds, the Waxing/Waning mode stimulus may appear as a very brief flicker, and not as a coherent grating. The highest speed proportion of the measured tuning surface might therefore be driven by a different process that is not related to the processing of motion, which might bias the tuning surface towards the higher speed, because flickers can elicit very robust responses. An initial 1 second period where the grating remained stationary was introduced to the Standard and Moving mode to remove the effects of this flicker response, but this design could not be used in the Waxing/Waning mode. Further research is needed to determine if this explains the shift in the preferred speed.

The presence of a large population of neurons in our sample showing only robust speed-tuning when stimuli were presented in the Moving mode is consistent with the idea that the visual system has evolved to be most efficient in encoding naturally occurring features (Barlow, 1961; Simoncelli & Olshausen, 2001). This idea has been tested by Nishimoto & Gallant (2011), who studied the receptive field structure of MT neurons in awake macaques viewing motion-enhanced natural movies.

In that study, three-dimensional spectral receptive fields were estimated by fitting the responses to a multi-filter model with static nonlinearity for individual filter outputs, which

were then normalised and linearly summed. A continuum of speed selectivity was found, with a significant population of neurons showing speed-tuning. However, since the 3-dimensional plan fitting procedure used for quantifying speed-tuning in Nishimoto and Gallant (2011) is quite different from the Q_{sp} parameter used in the present study (and analogous estimates in other previous studies), it is difficult to directly compare the proportion of speed-tuned neurons. Our results show that the integration of information across frequency channels in MT is dependent on the content of the stimulus, and can be changed from being non speed-tuned to speed-tuned by a manipulation as simple as displacing the Gabor pattern on a trajectory in time.

The data shown in Figure 3.7 indicate that there are incremental changes in the dependency between the SF of the stimulus and the optimal speed across the entire dataset, rather than suggesting the existence of a specific population of MT cells. Contextual effects associated with motion trajectories in the Moving mode might be mediated by propagation of signals via intrinsic horizontal connections in MT (Malach et al., 1997; Ahmed at al., 2012) or by feedback circuits originated from higher order visual areas, such as the medial superior temporal (MST) or the lateral intraparietal (LIP) areas.

Differences in the shapes of the SF tuning curves were minor in the three modes, suggesting that the representation of spatial frequency in MT is fairly invariant to the characteristics of the motion stimulus. Interestingly and unexpectedly, the optimal speed is shifted slightly toward higher speeds in the Moving mode in addition to detecting a larger population of speed-tuned neurons.

4. Representation of the speed of motion in the primary visual cortex

The primary visual cortex (V1) is the gateway to the primate visual cortex, and it is one of the major sources of inputs to area MT. An important question about the hierarchical organisation of the visual cortex is the extent to which response characteristics of MT neurons are inherited from the characteristics that are already present in V1. While early studies using the traditional drifting gratings have shown little-to-no speed tuning in V1, the discovery described in Chapter 3, showing that motion trajectories can modulate speed-tuning in MT, raises two questions: Is the contextual effect observed in MT already present in V1 and, has the population of speed-tuned V1 neurons been underestimated with drifting gratings? This chapter addresses these two questions. Parts of the results described in this chapter have been published in Davies et al. (2016) *Sci. Reports*.

4.1 Abstract

Despite the fact that the visual cortex is often conceptualised as a hierarchical system, where cortical areas at higher levels of the hierarchy perform more elaborate computations than those performed by lower-level cortical areas, the extent to which the response characteristics of neurons in extrastriate areas can be considered separate from those derived from lower levels, has been a controversial topic. Contrary to the expectation of hierarchical processing, in the domain of motion processing where area MT is known to play a central role, key tuning characteristics of MT neurons such as direction selectivity and speed-tuning have been argued to be derived from computations performed by V1 neurons in V1. In Chapter 3, it was demonstrated that motion trajectories – an important feature of moving patterns in the natural environment – can alter how motion speed is processed in area MT of anaesthetised marmosets, revealing a larger population of speed-tuned neurons than what have been previously reported. In this study, we examined the responses of V1 neurons, with the same protocol, in order to determine whether this phenomenon could be derived from V1 inputs, or are most likely resulting from MT computations. Motion trajectories were not found to increase the degree of speed-tuning of V1 cells to the same extent as in MT. In fact, only a small population of V1 neurons was found to be speed-tuned (6/87=7%), when tested with stimuli incorporating motion trajectories. This percentage was a slight increase from what was found with the traditional drifting grating stimulus (1/87=1.1%). The results suggest that the non-linear summation behaviour of motion signals in MT described in Chapter 3 is not likely to be derived solely from processing performed by V1 neurons.

4.2 Introduction

The visual cortex is commonly conceptualised as a hierarchical system where visual information passes through successive cortical areas, processing progressively more elaborate visual information. The V1-to-MT pathway of motion processing is often cited as a classic example of hierarchical processing. The fact that MT neurons have larger receptive fields than V1 neurons, and that a much larger proportion of MT neurons are direction selective in comparison to that of V1 (Section 1.2.2), suggest that MT is more specialised in motion processing at a larger spatial scale, relative to V1. However, the extent to which the behaviours of MT neurons represent novel computation, not already present in V1, remains an area where many questions are unresolved.

Direction selectivity, the defining feature of MT, has been proposed to be a characteristic that is intrinsic to MT. The theory is motivated by experiments which either lesioned or inactivated V1, reporting that most of the MT neurons that remained responsive were direction selective (Rodman et al., 1989; Girard et al., 1992; Rosa et al., 2000). The results suggest that direction selectivity in MT is not inherited from V1 (since it is deprived of its V1 inputs in these experiments), but is generated *de novo* by the intrinsic circuitry of MT. However, since the majority of MT neurons are actually silenced by V1 lesion or inactivation, it is possible that the theory is only valid for specialised populations of MT neurons, which receive driving inputs from subcortical sources (see Section 1.2.1), and does not apply to the majority of MT neurons.

An argument against direction selectivity being intrinsic to MT is that, despite the large receptive fields of MT neurons, the spatial extent in which the displacement of a bar stimulus can trigger direction selective responses is very small – in fact, comparable to the size of single V1 receptive fields (Mikami, 1992; Livingston et al., 2001). This observation suggests that the large MT receptive field might be aggregated from a large number of "subfields",

which are driven by V1 neurons that are already direction selective. The study of Movshon and Newsome (1996), provided more direct evidence for this idea by demonstrating that V1 neurons that directly project to MT are themselves already direction selective (Section 1.3).

The resolution of locally ambiguous motion directions (i.e., the aperture problem; Section 1.6.2) is another area that has been extensively studied. Studies using the "plaid" stimulus (the superimposition of two sinusoidal gratings drifting at two different directions) have reported that about 20% of MT neurons, called "pattern cells", are tuned to the global directions of the plaids, instead of to the directions of the individual gratings (Movshon et al., 1985; Smith et al., 2005; Solomon et al., 2011). This is in contrast to V1, where no such behaviour was found by Movshon et al. (1985). Movshon & Newsome (1996) showed that V1 neurons identified to directly project to MT were not pattern cells, further enforcing the idea that the disambiguation of the directions of plaids is a function that emerges from the V1-to-MT circuitry. However, small populations of pattern cells have been subsequently reported in V1 (~1% in Tinsley et al., 2003; 9% in Guo et al., 2004), which have called the theory into question. Whether the behaviour of MT pattern cells is inherited from a small population of V1 neurons that are already pattern selective remains a possibility.

Compared to the two characteristics discussed above, the coding of motion speed in V1 and MT is relatively under-studied. Priebe et al. (2006) examined the response patterns of V1 neurons in the macaque monkey using drifting sinusoidal gratings, and reported a significant population of speed-tuned, direction selective neurons in V1. The proportion of speed-tuned, direction selective neurons in V1 (~25%) was found to be comparable to that in MT (Priebe et al., 2003), which supports the idea that speed-tuning in MT might be primarily inherited from V1. However, this does not necessarily mean that the coding of motion speed in V1 and MT is identical. Priebe et al. (2003) demonstrated that in MT, Fourier components can

interact nonlinearly to give rise to more reliable speed-tuning (Section 1.6.4), but this characteristic was not observed in V1 (Priebe et al., 2006).

The results discussed in Chapter 3 demonstrated that the extent to which a MT neuron behaves in a speed-tuned manner is dependent on the stimulus that is used, and that motion signals associated with trajectories (the Moving mode stimulus) can reveal a higher proportion of speed-tuned neurons, which might be mediated by non-linear summation. These new findings naturally motivated experiments in V1 using identical protocol, in order to re-evaluate the relationship between V1 and MT in terms of speed-tuning.

4.3 Methods

Single-unit recordings were obtained from area V1 of 3 adult New World monkeys (*Callithrix jacchus*, the common marmoset). Experiments were conducted in accordance with the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes, and all procedures were approved by the Monash University Animal Ethics Experimentation Committee.

Animal preparation

Preparation of the animal for *in vivo* recordings of neurons was based on the protocol described in detail by Bourne and Rosa (2003), updated according to Yu and Rosa (2010). For more detail, see Section 2.1.

Quantitative tests

Electrophysiological recordings were used to test neurons sampled at between 100-300 μm intervals. At each location the boundary of the multiunit minimum response field was mapped using a manually operated stimulus. After receptive field isolation, the neuron was tested for its preferred orientation (and direction if the neuron was direction selective). After these initial tests, the three conditions of Standard, Moving and Waxing/Waning modes of moving stimulus were displayed. For direction selective neurons, the motion direction was set to be the preferred direction. For non-direction selective neurons, the direction of motion was set to be orthogonal to one of the preferred orientations. Each motion mode was presented in a 5x5 matrix of conditions, where the spatial frequency and the speed of motion were manipulated independently. For more detail on stimuli, see Section 2.2.

Data Analysis

To determine if a neuron is speed-tuned, the mean firing rate for each of the 5x5 conditions was calculated within a temporal window, which was set to capture the time interval where the Gabor pattern was inside the receptive field. This process produces a "tuning surface", which was fitted to two completion models, one of which assumes that the optimal speed does not vary with the spatial frequency of the stimulus, while the second allows for systematic relationship of speed and spatial frequency. The Bayesian Information Criterion was used to select the model that was more likely to be the better model for explaining the data. These models also estimated the optimal spatial frequency, the optimal speed, and other parameters of the tuning surfaces. Recordings were confirmed in primary visual cortex via histological reconstruction. For a more in depth description, including equations, see Section 2.3.

4.4 Results

We characterized the response properties of 87 neurons, confirmed to be in V1 by histological reconstruction. The receptive fields were in the near periphery, between 10^o and 30^o eccentricity, encompassing both the upper and lower contralateral quadrants. This level of eccentricity was chosen to match a similar population to that which was measured in the previous chapter, to allow for better comparison.

The responses of neurons fell into two main categories. The first of these was a temporal frequency-tuned response profile for all three conditions, as illustrated in Figure 4.1. The responses are illustrated as space-equalised peristimulus time histograms (PSTHs, left) and speed-tuning curves obtained with stimuli of different spatial frequencies (centre). The PSTHs were converted into tuning surfaces which were then fitted with two models (right): an SF-speed separable model (M_0), which assumed that SF and speed-tuning curves were independent of each other, and an SF-speed inseparable model (M_1) in which the optimal speed could vary systematically with the stimulus SF. We then use the Bayesian Information Criterion (BIC; Yu et al., 2010) to calculate $P(M_0|M_0,M_1)$, the probability that M_0 provided a better fit to the data than M_1. For this neuron, $P(M_0|M_0,M_1)$ was lower than 0.001 for stimulation in all three modes, indicating that it was not speed-tuned, regardless of the stimulus used.

Figure 4.2 shows the response pattern of one representative neuron from a small population (n = 4), for which the response only in the Moving mode was speed-tuned. For this neuron, $P(M_0|M_0,M_1)$ was lower than 0.001 for both of the drifting gratings, however for the moving Gabor patch $P(M_0|M_0,M_1)$ was 0.92, indicating that the pattern of responses was better described by a model where the neuron's optimal speed is independent of the spatial frequency of the stimulus.

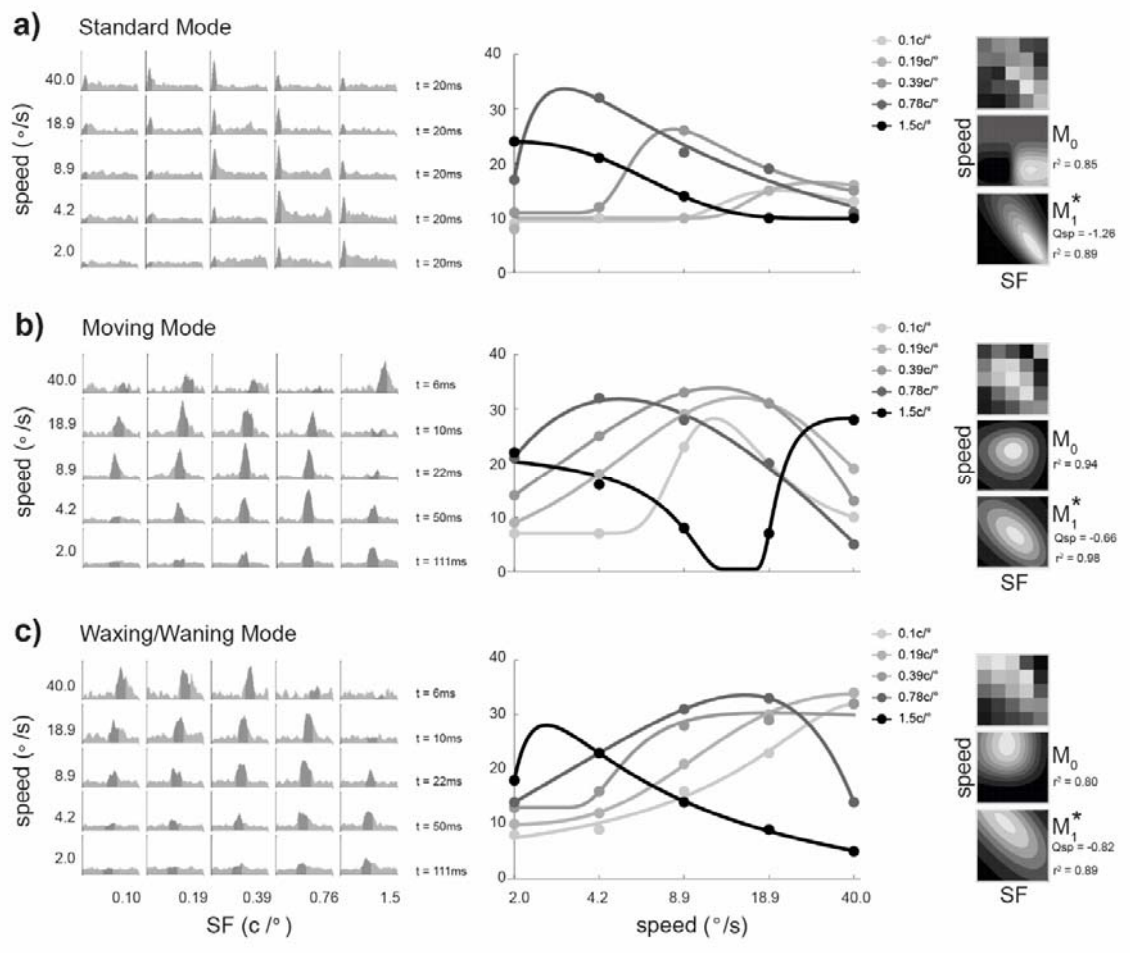

Figure 4.1. Data from a V1 neuron that was not speed-tuned in any presentation mode. Rows show responses to the a) Standard, b) Moving and c) Waxing/Waning modes. For each panel, the first column shows the responses to each SF-speed combination as a peri-stimulus time histogram (PSTH). In these histograms, bin sizes (indicated at the right hand side of each row) were adjusted to represent equal distances travelled by the stimulus in the Moving mode. The y-axis represents the response-rate calculated within the bins; scale is the same for all histograms, with the maximal value of 265Hz. The shaded regions represent the time intervals during which the response rates shown in the second and the third column were calculated. The second column displays speed-tuning curves obtained using Gabor patches of different SF. The circles represent the measured response rates, and the curves are fitted skewed log-Gaussian functions. The third column illustrates the tuning surface describing the responses to the 5 x 5 matrices of conditions (top) and the result when surfaces were fitted to two models, M_0 (middle) and M_1 (bottom). The models that provided the better fit, as determined by $P(M_0|M_0,M_1)$are indicated by asterisks. The quality of fit (r^2) and the estimated Q_{sp} are also provided.

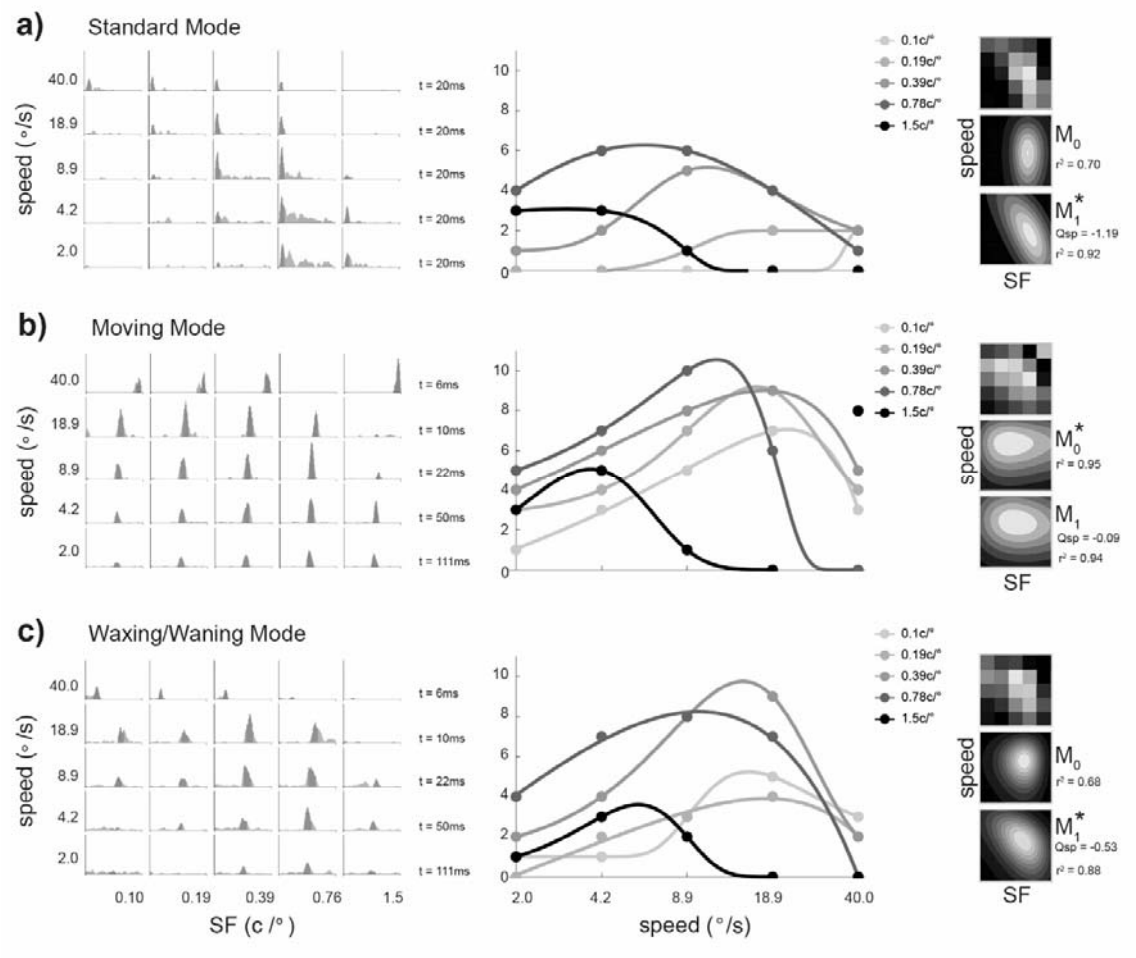

Figure 4.2. Data from a V1 neuron that was speed-tuned only in the Moving mode. The data are presented in the same format as Figure 4.1. The y-axis in all PSTHs uses the same scale (maximum 130Hz).

For the neuron illustrated in Figure 4.1, the Q_{sp} parameter (Fig. 2.6) estimated from the fitted model (third column) for the Standard, Moving and Waxing/Waning modes were -1.26, -0.65, and -0.82 respectively. The confidence intervals for each of these values included -1, but did not include 0, indicating that the neuron was TF-tuned. For the neuron illustrated in Figure 4.2, Q_{sp} were -1.20, -0.08 and -0.53 respectively. Confidence intervals indicated a temporal frequency-tuned response for the Standard mode, where the confidence interval was inclusive of -1 but not inclusive of 0; the Moving mode was inclusive only of 0, indicating speed-tuning. The confidence interval for the Waxing/Waning mode did not include either 0 or 1, indicating that the speed-SF dependency was intermediate between being perfectly speed-tuned and a perfectly TF-tuned.

4.4.1 Speed-tuning classification across stimulus modes

Using the $P(M_0|M_0,M_1)>0.5$ criterion, our main finding is that for the majority of the V1 neurons (81 neurons; 94%), there was no effect of presentation mode (in comparison to 52.7% in MT; Table 3.1). These neurons were consistently non-speed tuned. However, there was a small population of neurons (6 neurons; 7%) that were found to be speed-tuned under the Moving mode (in comparison to 48.2% in MT). Of these six neurons, one was also classified as speed-tuned in the Standard mode, and another was also classified as speed-tuned in the Waxing/Waning mode (See Table 4.1). The median $P(M_0|M_0,M_1)$ values for the Standard, Moving, and Waxing/Waning mode were 0.00000027, 0.11, and 0.007 respectively (Figure 4.3, right column). These values were significantly different (Wilcoxon Test; Standard vs Moving: $V = 735$, $p<0.001$; Standard vs Waxing/Waning: $V = 1175$, $p<0.001$; Moving vs Waxing/Waning: $V = 3174$, $p<0.001$).

The left column of Figure 4.3 plots the distribution of Q_{sp}. In the Standard mode, the median value for Q_{sp} was -0.93, whereas in the Moving and Waxing/Waning modes, the median value was shifted to -0.54 and -0.58 respectively (Wilcoxon Test; Standard vs Moving: V =

3799, p<0.001; Standard vs Waxing/Waning: V = 910, p<0.001). While the median values for the Moving and the Waxing/Waning mode were comparable, the difference was significant (Wilcoxon; V = 2674, p = 0.02).

Table 4.1. Speed-tuning classification of neurons across the three test conditions. Speed-tuning was determined using a threshold criterion of $P(M_0|M_0,M_1) > 0.5$.

Standard Mode	Moving Mode	Waxing/Waning Mode	Number
Speed-Tuned	Speed-Tuned	Speed-Tuned	0 (0%)
Speed-Tuned	Non Speed-Tuned	Speed-Tuned	0 (0%)
Speed-Tuned	Speed-Tuned	Non Speed-Tuned	1 (1.1%)
Speed-Tuned	Non Speed-Tuned	Non Speed-Tuned	0 (0%)
Non Speed-Tuned	Speed-Tuned	Speed-Tuned	1 (1.1%)
Non Speed-Tuned	Non Speed-Tuned	Speed-Tuned	0 (0%)
Non Speed-Tuned	Speed-Tuned	Non Speed-Tuned	4 (4.6%)
Non Speed-Tuned	Non Speed-Tuned	Non Speed-Tuned	81 (9.3%)

4.4.2 Speed-tuning characteristics across direction selective V1 neurons

Of the 87 V1 neurons tested, 13 neurons were found to be direction selective. None of the direction selective neurons were found to be speed-tuned under any of the conditions.

Median values of $P(M_0|M_0,M_1)$ were 0.0, 0.19 and 0.03 for the Standard, Moving and Waxing/Waning modes respectively. There was no significant difference between the Waxing/Waning mode and either the Standard (Wilcoxon test; V = 23, p = 0.1272) or the Moving mode (Wilcoxon test; V = 68, p = 0.1272). There was however a significant different between the Standard and the Moving modes (Wilcoxon test; V = 6, p = 0.003).

Median values for Q_{sp} were -0.72, -0.64 and -0.54 for the Standard, Moving and Waxing/Waning modes respectively. However, none of these differences were found to be significant (Wilcoxon test; Standard vs Moving: V = 21, p = 0.09; Standard vs Waxing/Waning: V = 29, p = 0.27; Moving vs Waxing/Waning: V = 36, p = 0.8445.)

In Figure 4.3, the distributions of $P(M_0|M_0,M_1)$ and Q_{sp} of direction selective neurons (shaded in a lighter grey) are compared against those of non-direction selective neurons (shaded in a darker grey). The distributions of the two populations were similar, particularly for the

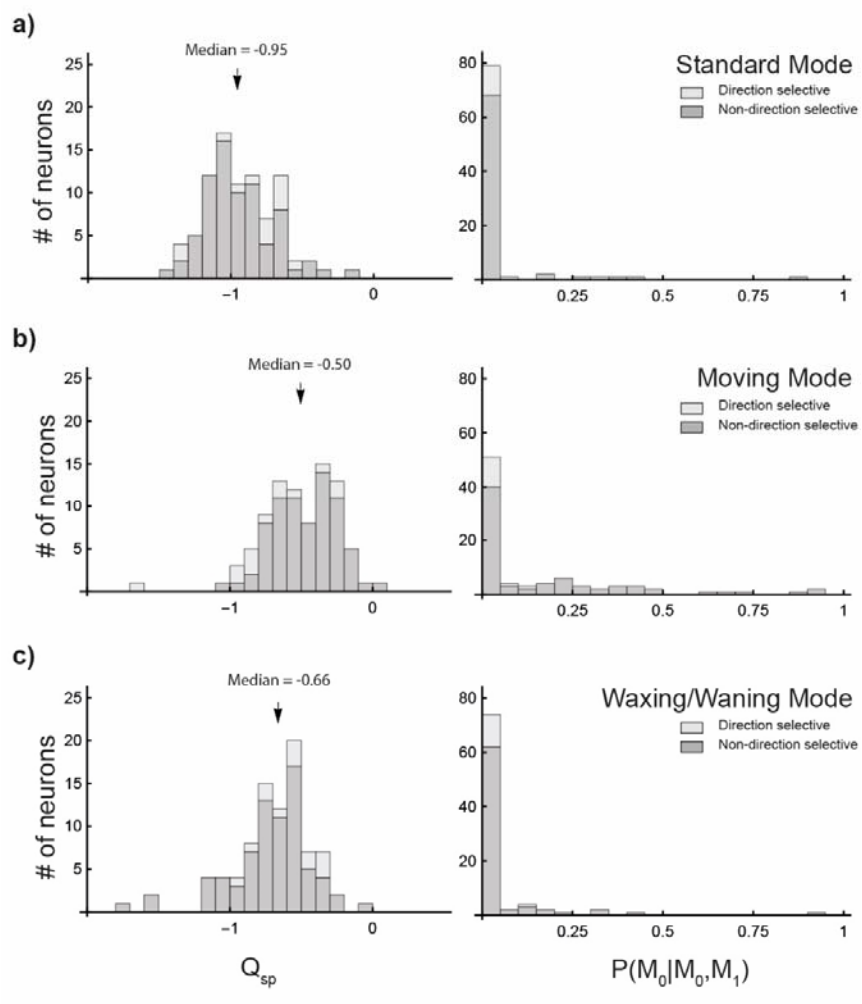

Figure 4.3. Population graphs showing the distribution of Q_{sp} and $P(M_0|M_0,M_1)$. Direction selective neurons are indicated by the lighter grey shading, all other neurons are indicated by the darker shade of grey. Black arrows on the Q_{sp} graph indicate population median values.

Standard (Mann-Whitney; Q_{sp}: U=436.5, p=0.11, $P(M_0|M_0,M_1)$: U=519.5, p=0.42) and the Waxing/Waning modes (Mann-Whitney; Q_{sp}: U=454.5, p=0.15, $P(M_0|M_0,M_1)$: U=644.5, p=0.70). There was however a small shift of 0.08 for non-direction selective neurons in $P(M_0|M_0,M_1)$ in the Moving mode when compared the direction selective neurons (Mann-Whitney U; U = 828.5, p = 0.031).

4.4.3 Spatiotemporal tuning characteristics of V1 neurons

After classification of speed-tuning, the better model of the two was selected according to $P(M_0|M_0,M_1)$. Within both models were eight parameters (M_1 had a ninth parameter of Q_{sp}) that quantified the spatiotemporal response characteristics of the neurons (for equations see Chapter 2). These parameters included the scaling parameters *b* and *c*, which quantify the range of the firing rate of the neuron. As determined by a two-way ANOVA, there was no significant effect of stimulus mode on these two parameters (F(1,92) = 1.090, p = 0.3086), suggesting that the firing rate of the neuron is not affected by the stimulus mode.

Three parameters determine the neuron's spatial frequency tuning curve (see Section 2.3.1 for further information): 2^{sf_0} represents the optimal spatial frequency, σ_{sf} characterises the bandwidth of the tuning curve, and ζ_{sf} the skewness. The median values of 2^{sf_0} were 0.39 c/°, 0.33 c/° and 0.29 c/° for the Standard, Moving and Waxing/Waning modes respectively (Figure 4.4, left column). Pairwise comparisons highlighted significant differences between all three modes. The Standard mode was found to have a significantly higher optimal spatial frequency when compared to the Moving (Wilcoxon; V = 3106, p<0.0001) and the Waxing/Waning modes (Wilcoxon: V = 3589, p<0.0001). The Waxing/Waning mode was found to have the lowest spatial frequency, also differing significantly from the Moving mode (Wilcoxon: V = 2891, p = 0.007).

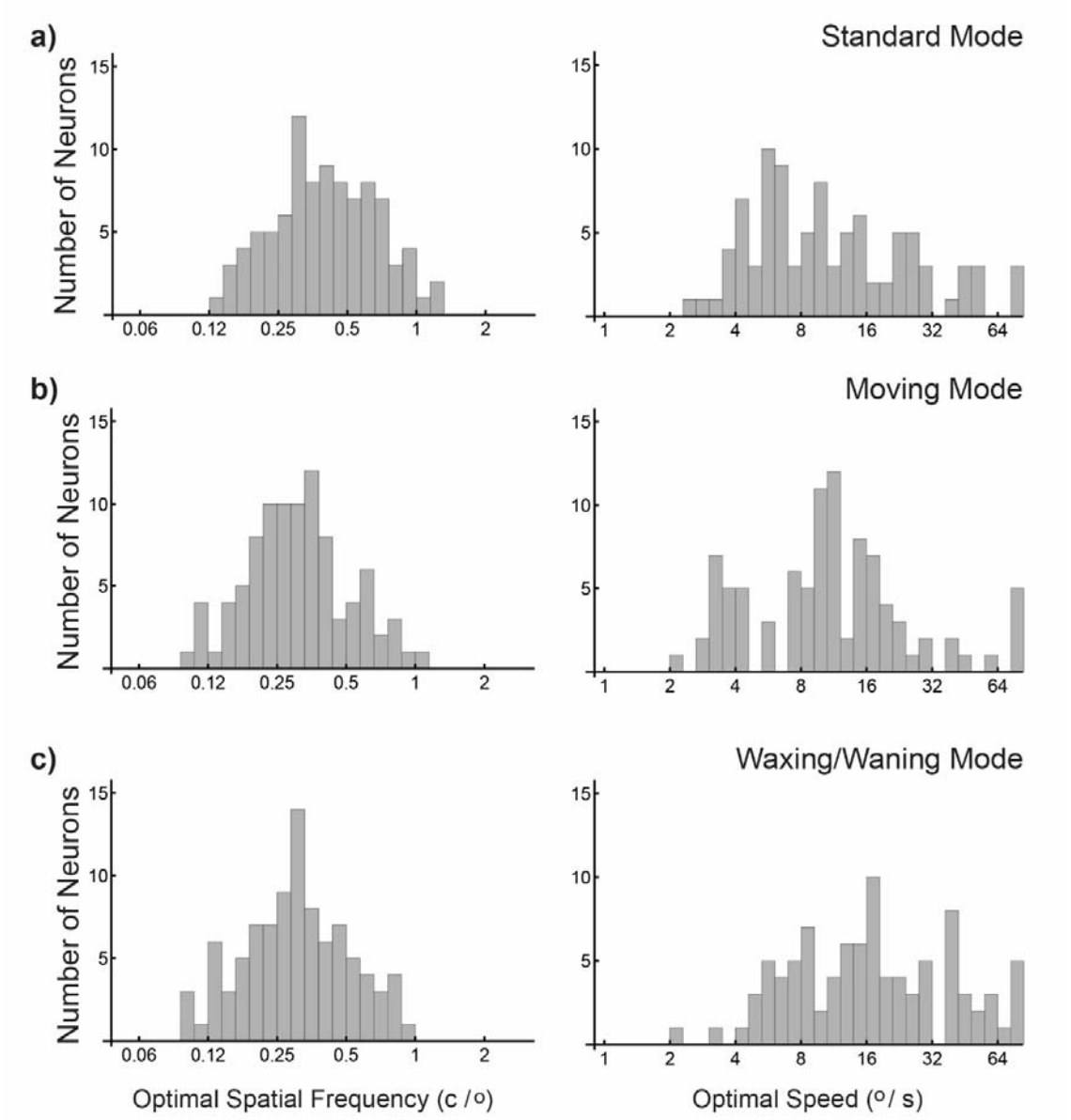

Figure 4.4. Population graphs showing the distribution of the optimal spatial frequency (left column) and speed (right column) as calculated by the chosen model. Black arrows indicate the median value.

For σ_{sf}, median values were 1.5, 1.52 and 1.58 for the Standard, Moving and Waxing/Waning mode respectively. None of these values were found to be significantly different from each other when tested with pairwise comparisons. For the ζ_{sf} skewness variable, median values were 0.06, 0.008, and 0.005 for each of the three modes respectively. Pairwise comparisons showed the Standard mode was found to be significantly different from both the Moving

(Wilcoxon: V = 1085, p = 0.003) and Waxing/Waning modes (Wilcoxon: V = 835, p = 0.001). However there was no significant difference between the Moving and Waxing/Waning modes (Wilcoxon: V = 990.5, p = 0.74).

Similar to the spatial frequency tuning parameters, three parameters determined a neuron's speed-tuning curve: 2^{sp_0} represents the neurons optimal speed, while σ_{sp} characterises the bandwidth of the tuning curve, and ζ_{sp} the skewness. The median values of 2^{sp_0} were 9.91 °/s, 10.95 °/s and 18.61 °/s for the Standard, Moving and Waxing/Waning modes respectively (Figure 4.4, right column). Pairwise comparisons highlighted significant differences between all three modes. The Waxing/Waning mode was found to have a significantly higher optimal speeds when compared to the Standard (Wilcoxon; V = 513, p<0.0001) and the Moving modes (Wilcoxon: V = 481, p<0.0001). The Standard mode was found to have the lowest optimal speed, however this was not significantly different from the optimal speed for the Moving mode (Wilcoxon: V = 2193, p = 0.69).

For σ_{sp} , median values were 1.66, 1.42 and 1.63 for the Standard, Moving and Waxing/Waning mode respectively. The Moving mode was significantly different from both the Standard (Wilcoxon: V = 2583, p = 0.030) and the Waxing/Waning modes (Wilcoxon: V = 1530, p = 0.025), however there was no significant difference between these two modes (Wilcoxon: V = 2067, p = 0.65). For the ζ_{sp} skewness variable, median values were 0.003, 0.08, and 0.04 for each of the three modes respectively. Pairwise comparisons showed the Standard mode was found to be significantly difference from both the Moving (Wilcoxon: V = 1013, p = 0.002) and Waxing/Waning modes (Wilcoxon: V = 706, p = 0.0002). However there was no significant difference between the Moving and Waxing/Waning modes (Wilcoxon: V = 1139.5, p = 0.25).

4.5 Discussion

Since cortical area V1 is the largest source of input to area MT (Section 1.3), the characteristics of speed processing in V1 is a critical issue in understanding the hierarchical structure of motion processing in the primate visual cortex. The relative proportion of speed-tuned neurons in the two areas is of particular interest. The issue has been previously investigated in the macaque monkey by Priebe et al. (2003, 2006). Using drifting gratings (similar to our Standard mode stimulus), they reported that the populations of speed-tuned neurons in macaque MT and V1 were small and comparable (~25% in both areas). In comparison, we also found that the proportions of speed-tuned neurons in MT and V1 were similar using our Standard mode stimulus, except that the percentage was significantly smaller (~1% in both areas). Our results are therefore consistent with Priebe et al. (2003, 2006), suggesting that MT does not have significantly more speed-tuned neurons (when tested with the traditional drifting gratings), relative to V1. This is in dramatic contrast to the population of direction selective neurons, which represent about 20~30% of the neurons in V1 (Yu & Rosa., 2014), but over 85% of the neuron in MT (Section 1.2.2).

The differences in the percentage of speed-tuned cells in the marmoset (~1%) and in the macaque (~25%) might reflect a genuine species difference. However, given that close similarities in response characteristics in V1 and MT in the two species have been reported (Section 1.7), other factors must be considered. One possible explanation is the stimulus that was used. The Standard mode stimulus used Gabor patterns, which avoided the sharp boundaries in the traditional drifting grating patches used in Priebe et al. (2003; 2006). Given the nonlinear summation behaviour that was reported in Chapter 3 (as well as those by Priebe et al., 2003; 2006), the broadband spatial frequencies introduced by the sharp boundaries are expected to increase the number of speed-tuned cells. It is therefore likely that the observed difference is attributed simply to the Fourier contents of the stimuli.

An additional factor that is worthy of consideration is the location of the receptive fields: whereas Priebe et al. (2006) studied neurons with receptive fields in the central visual field, our receptive fields were in the near peripheral visual field (10°-30° in eccentricity). We consider this an unlikely explanation, because although the processing of motion signal is very often believed to be dependent on eccentricity, experimental evidence has been reported to refute this (Yu et al., 2015). In addition, Yu et al. (2010) quantified speed-tuning in the marmoset V1 across a large extent of the visual field, and was unable to find significant differences.

There is one caveat in the comparison that we made above: In the experiment by Priebe et al. (2006), only direction selective V1 neurons were sampled, whereas in our study, we sampled 13 direction selective V1 neurons and 74 non-direction selective neurons in V1. Although in Section 4.4.2, we concentrated on the direction selective neurons in our sample, and arrived at the same conclusion, the small sample size of direction selective neurons made direct comparison to the results of Priebe et al. (2006) difficult. Further studies, concentrated on characterising the speed-tuning characteristics of direction selective V1 neurons in the marmoset, are needed.

Another finding by Priebe et al. (2003) was the presence of nonlinear processing mechanisms in MT, which was demonstrated with the superimposition of two drifting gratings of different central spatial frequencies drifting at the same speed. This type of stimulus was found to increase the extent to which speed tuning was invariant to spatial frequency in area MT (Priebe et al., 2003), but the effect was not present in V1 (Priebe et al., 2006). The result suggested that although the speed-tuning in macaque MT may be derived from V1 inputs, MT neurons have additional nonlinear summation characteristics which can lead to more robust speed-tuning, given more complex stimuli.

In Chapter 3, we demonstrated a different form of nonlinear summation in the marmoset MT; with the Moving mode stimulus, resulting in the proportion of speed-tuned neurons in MT was increased from 0.9% to 48.2%. The data presented in this chapter showed that in V1, the effect was much weaker: the number of speed-tuned neurons was increased from 1.1% (Standard mode) to 6.8% (Moving mode). The result is therefore consistent to Priebe et al. (2006).

In summary, Priebe et al. (2003, 2006) concluded that when tested with drifting gratings, moderate percentages (~25%) of neurons in macaque V1 and MT were found to be speed-tuned, and that speed-tuning of MT neurons was likely to be inherited from V1 neurons. In contrast, our data showed that in marmoset V1 and MT, the percentages of speed-tuned neurons are exceedingly small (~1%). In addition, we showed that when tested with the Moving mode stimulus, about half of the neurons in MT are speed-tuned, whereas in V1, speed-tuned neurons remain rare. Therefore, we conclude that speed tuning in marmoset MT is most effective in a naturalistic setting, when contextual information such as motion trajectories are present. In such conditions, speed-tuning in MT is not inherited from V1 inputs, but is the result of nonlinear processing that operates primarily in MT. Further studies which sample more direction-selective V1 neurons are needed to further substantiate this claim.

5. Response patterns to Moving mode stimulus in simulated MT neurons

In Chapter 3, I reported that the population of speed-tuned neurons in cortical area MT is significantly larger than previously reported, but only if the Moving mode stimulus was used to test speed-tuning. I argued informally that the finding is unexpected according to the linear pooling theory of information integration in MT (Section 1.6.4), and therefore suggests a nonlinear mechanism that has not been demonstrated before. In this chapter, I represent a formal version of the argument, using computational modelling to demonstrate quantitatively that the experimental results call for a re-evaluation of the linear pooling theory.

5.1 Abstract

While several important features about the integration of motion information at the level of cortical area MT have been successful explained by models that pool input signals linearly (Simoncelli & Heeger, 1998; Rust et al., 2006; Bradley & Goyal, 2008; Nishimoto & Gallant, 2011), recent studies have reported non-linear summation behaviours for MT neurons, particularly in the domain of the representation of speed (Priebe et al., 2003; Davies et al, 2016). In particular, Davies et al. (2016; see chapter 3) demonstrated that the extent to which MT neurons encode the speed of motion signals in a spatial frequency-invariant manner was dependent on the stimulus. While the majority of MT neurons did not show speed-tuning when tested with the conventional drifting grating stimulus, a signification population was found to be speed-tuned, when tested with a sinusoidal pattern that traversed the visual space in a trajectory. To examine if the results of Davies et al. (2016) was a novel phenomenon, not predicted by the linear-pooling model, we constructed models of MT neurons to fit the data of Davies et al. (2016), obtained with the conventional drifting grating stimulus. The models followed the general scheme of Simoncelli & Heeger (1998), where the outputs of motion-energy filters were normalised and linearly pooled, and were used to predict the responses to the patterns moving in trajectories ("Moving mode"). Of all the 110 MT neurons reported in Davies et al. (2016), only 1 was predicted to change from being non speed-tuned when tested with the conventional stimulus to being speed-tuned when tested with the Moving mode stimulus. Except for one additional case, where the model was speed-tuned under both stimulus modes (which faithfully duplicated the experimental results), the rest of the models remained non speed-tuned under Moving mode. The models were not able to duplicate the 45.5% of the MT neurons reported in Davies et al. (2016), which became speed-tuned in the Moving mode. Further analyses revealed that the Moving mode stimulus tended to reduce the dependency of speed-tuning with respect to the spatial frequency contents of the stimulus, but

the reduction was not sufficient to turn the model neurons to be classified as speed-tuned. The linear pooling model that was commonly used to explain the behaviours of MT neurons was therefore unable to explain the results of Davies et al. (2016).

5.2 Introduction

In Chapter 3, we reported that for a significant population of MT neurons (47.3%), speed-tuning was dependent on the stimulus used: they exhibited the classic temporal frequency-tuned behaviour when tested with the conventional drifting grating stimulus (Standard mode), but when they were tested with the Moving mode stimulus, they became speed-tuned. We informally argued that the result suggested a novel phenomenon, unexpected by a prominent class of theories of MT computation. The argument can be summarised as the following: the most successful model for information integration at the level of MT posits that inputs from V1 are pooled by a linear summation mechanism (Section 1.6.2). Given that the Standard mode stimuli consist of narrow ranges of spatiotemporal frequencies (Fig. 2.4a), the tuning surface obtained with the Standard mode stimulus therefore constitutes an estimate of the summation weights. Assuming linear summation, the tuning surface of the neuron when it is tested by the Moving mode stimulus is predictable: it is just the product of the Fourier representation of each condition in the Moving mode (Fig. 2.4b), and the tuning surface (as estimated by the Standard mode stimulus). The result should remain identical to the tuning surface estimated by the Standard mode stimulus. A temporal frequency-tuned neuron, tested with the Standard mode stimulus, therefore should remain temporal frequency-tuned when tested with the Moving mode stimulus. MT neurons that became speed-tuned under Moving mode therefore cannot be explained by models based on linear summation.

In this chapter, the qualitative argument outlined above is developed into a quantitative one with computational modelling. Quantitative modelling was needed to complement the qualitative argument for a couple of reasons: 1. Although linear summation is at the heart of the theories that we are addressing, nonlinear components (such as energy filtering and normalisation; see Section 1.6.2) are present in those models. It is possible that these nonlinear elements might change the shape of the tuning surface when tested with the

Moving mode. 2. In the qualitative argument, the tuning surface and the Fourier representations of the stimulus are conceptualised as continuous functions, while in the physiological experiments, they are sampled by a small number (5x5) of discrete combinations of spatial and temporal frequencies. The surface fitting procedure might bias the estimated shape of the tuning surface in unexpected ways. 3. The Bayesian procedure (Section 2.3.1) used to determine if a neuron is speed-tuned is nonlinear. The first factor attributes the findings to the nonlinear components in current models of MT neurons, while the second and the third factors attribute the finding to nonlinearity in the analysis procedure. They might operate by themselves or interact with each other to change speed-tuning classification in the Moving mode. Because the dependency between speed-tuning and the SF of the stimulus forms a continuous distribution (Fig. 3.5b), small differences in the Q_{sp} parameter might be amplified by these nonlinear operations, which might turn a non-speed tuned neuron into a speed-tuned neuron.

These factors are addressed in this chapter by the following modelling approach: A computational model for individual MT neurons, following the general scheme outlined in Section 1.6 was constructed for each MT neuron studied in Chapter 3. Each model MT neuron pooled the outputs of a population of V1 neurons (Section 5.3.2), whose weights were determined by fitting the output of the model to the tuning surface obtained with the Standard mode stimulus. The model was then used to predict the responses to the Moving mode stimulus. If the result was similar to the electrophysiological data, it was concluded that the phenomenon reported in Chapter 3 was the expected outcome given current knowledge about MT. If the result was not similar to the data, it was concluded that the phenomenon was novel, unexpected by the linear pooling theory of MT processing.

5.3 Methods

For each of the MT neurons studied in the electrophysiological experiments (reported in Chapter 3), we constructed a computational model that was fitted to the experimental data obtained with the Standard mode stimulus. The model was based on the linear pooling of the responses of a population of simulated V1 neurons. The model's responses to the Moving mode stimulus were then compared to those obtained in the electrophysiological experiment.

The spatial patterns used in the electrophysiological experiments described in Chapter 3 were two-dimensional in space, but since the luminance of the spatial pattern varied only in one dimension (which was always paralleled to the direction of motion; see Section 1.4.2), the stimuli were effectively one-dimensional in space (Fig. 5.1). In this chapter, all the simulations were therefore conducted on a two-dimensional stimulus space – 1 dimension for space and 1 dimension for time.

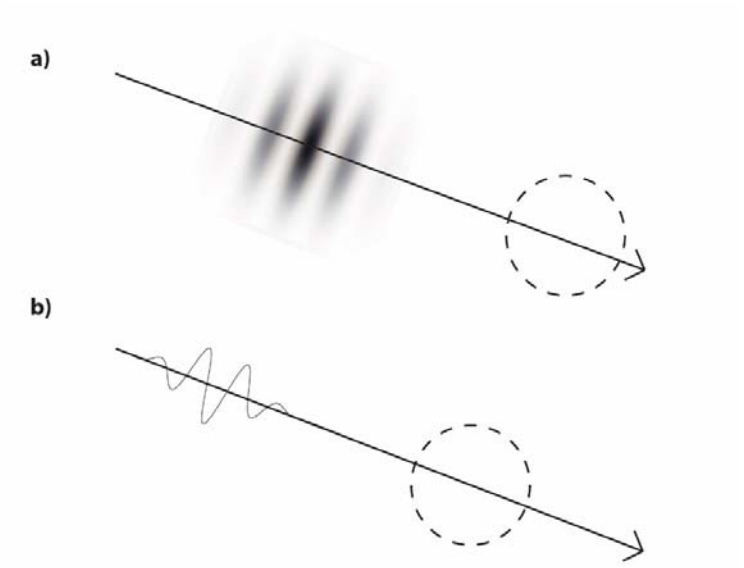

Figure 5.1. The Gabor pattern used in the Moving mode (the same principle applies to the Standard mode stimulus as well). a) A 2D Gabor pattern that moves in a trajectory across the receptive field (indicated by dashed circle). The direction is paralleled to the preferred direction of the receptive field. b) In the simulation, the Gabor pattern is represented by its profile, reducing the spatial dimension from 2 to 1.

5.3.1 Modelling V1 neurons with Gabor filters

The tuning characteristics of individual V1 neurons were modelled by spatiotemporal Gabor filters. As Gabor filters capture several prominent tuning properties found in V1 neurons (Jones & Palmer, 1987; Daugman, 1985; Movellan, 2002), they are commonly used for modelling spatiotemporal coding in V1.

Gabor filters have the mathematical form

$$V1_r(x,t) = e^{-\pi(a_{sf}*x)^2} * e^{-\pi(a_{tf}*t)^2} * e^{i2\pi*(sf0*x+tf0*t)} \quad (5.1)$$

The first exponential function represents a Gaussian envelope in space, the second a Gaussian envelope in time, and the third a drifting sinusoidal carrier. Due to the Euler's formula in the third exponential function

$$e^{ix} = \cos x + i \sin x \quad (5.2)$$

the complex function $V1_r(x,t)$ represents a pair of Gabor filters, whose carriers modulate according to a cosine function (the real part) and a sine function (the imaginary part). These two 90°-out-of-phase filters are called *quadrature pairs*, and they are used for calculating motion energy (Section 1.6.1, 5.3.3).

Since the carrier wave drifts in one direction (determined by the sign of tf_0), the Gabor filters used were all direction selective. In addition, since Equation 5.1 is the product of a spatial tuning function ($e^{-\pi(a_{sf}*x)^2} * e^{i2\pi*(sf0*x)}$) and a temporal tuning function ($e^{-\pi(a_{tf}*t)^2} * e^{i2\pi*(tf0*t)}$), it is spatiotemporally separable (Section 1.4.1), and therefore temporal frequency tuned rather than speed-tuned.

The a_{sf} parameter of the first term determines the size of the receptive field, the a_{tf} parameter of the second term determines the temporal window in which visual information is

integrated, and the sf_0 and tf_0 parameters of the third term control the optimal spatial and temporal frequencies of the neuron.

In the Fourier space, complex Gabor functions are bivariate Gaussian functions displaced from the origin (Daugman, 1985), indicating that they are bandpass filters in space and in time. If only the real part or the imaginary part of the Gabor function is used, in the Fourier space it is represented by a pair of bivariate Gaussian functions symmetrically displaced from the origin (Fig. 5.2, right col.).

Due to the inverse relationship between the size of the Gaussian envelope in the image space, and the size of the corresponding Gaussian distribution in the Fourier space (Daugman, 1985), the parameter a_{sf} also determines the bandwidth of the spatial frequency tuning curve (the larger the receptive field, the smaller the bandwidth for any one spatial frequency). Likewise, the parameter a_{tf} also determines the bandwidth of the temporal frequency tuning curve (the larger the window, the smaller the bandwidth).

The relationship is represented visually in Figure 5.2. In the left column, two receptive fields are illustrated in the image space: the one shown in Fig. 5.2a has higher central spatial frequency, smaller receptive field, higher temporal frequency, and shorter temporal integration interval than the one shown in Fig. 5.2b. In the Fourier space (right column), each filter is represented by a pair of Gaussian distributions, symmetrically displaced from the origin. Since the receptive field in Fig. 5.2a has higher optimal spatial and temporal frequencies, the centres of the Gaussian distributions are more removed from the origin, in both the spatial frequency (fx) and the temporal frequency axis (ft). The larger receptive field size of the unit shown in Fig. 5.2b is mirrored by the more concentrated Gaussian distributions in the spatial frequency axis (in other words, the receptive field has a sharper spatial frequency tuning curve). Likewise, the longer integration time is mirrored by the more

concentrated Gaussian distribution in the temporal frequency axis (in other words, the receptive field has a sharper temporal frequency tuning curve).

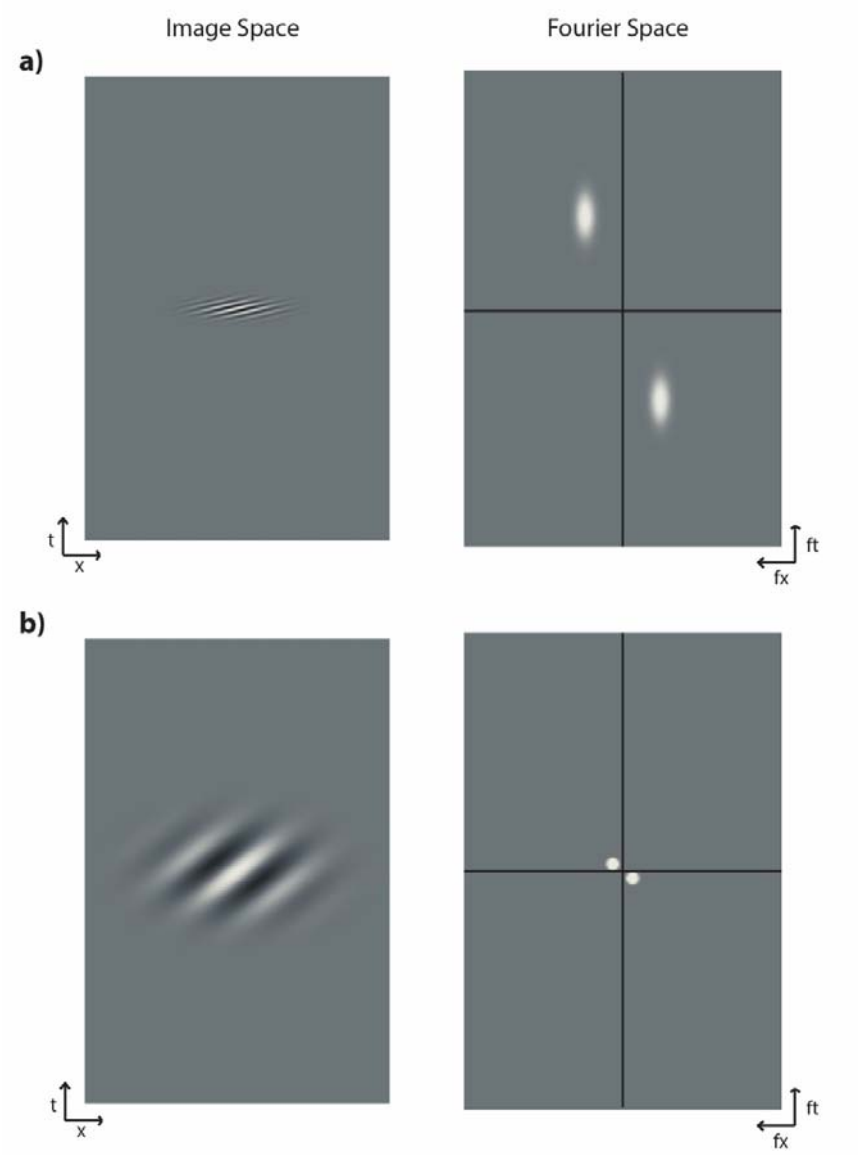

Figure 5.2. Gabor filters in the image space (left column) and in the frequency space (right column). In the left column, the axes are space (x) and time (t). In the right column, the axes are spatial frequency (fx) and temporal frequency (ft). The receptive field in a) has a higher optimal spatial frequency, higher optimal temporal frequency, smaller receptive field and shorter integration interval.

5.3.2 Modelling V1 neuron population

To sample the two-dimensional stimulus space (one for space and one for time), we used a population of 49 model V1 neurons, whose optimal spatial and temporal frequencies were uniformly sampled on the logarithmic scale. There was a quadrature pair of filters for each combination of 7 optimal spatial frequencies and 7 optimal temporal frequencies, which resulted in 98 filters (7x7x2). For each model MT neuron, these 98 filters sampled the same range of spatial and temporal frequencies that was sampled in the experiment.

For V1 neurons in primates and other mammals, the bandwidth (measured in octave) of spatial frequency tuning is fairly constant as a function of optimal spatial frequency. However, there is a small but systematic deviation from perfect constancy (Movshon et al., 1978; Tolhurst & Thompson, 1981; De Valois et al., 1982; Kulikowski & Vidvasagar, 1986; Yu et al., 2010). To reflect this inverse relationship, the half-amplitude bandwidth (in octave) of the spatial frequency tuning curve was set to be

$$2^{log_2 1.4 - 0.45 * log_2 \Delta sf * n} \tag{5.3}$$

where $n = 0 \ldots 6$ represents the 7 optimal spatial frequency that was sampled and Δsf represents a stepwise decrease of bandwidth as spatial frequency increases. Δsf was set to allow for the 7 values to be evenly distributed in log space from the defined minimum to maximum.

The temporal frequency was varied in a similar manner. To reflect this relationship, the half-amplitude bandwidth (in octave) of the temporal frequency tuning curve was set to

$$\sigma_{tf} = 2^{Log_2(2.8) - (0.35 * Log_2 \Delta tf * n)} \tag{5.4}$$

where $n = 0 \ldots 6$ represents the 7 peak temporal frequencies that were sampled and Δtf represents the required stepwise decrease of bandwidth as temporal frequency increases. Δtf

was set to allow for the 7 values to be evenly distributed in log space from the defined minimum to maximum.

The relationship between the σ_f parameter and the half-amplitude bandwidth is then represented by:

$$bandwidth = \frac{\ln \frac{\pi f_0 + \sigma_f \sqrt{\pi \ln 2}}{\pi f_0 - \sigma_f \sqrt{\pi \ln 2}}}{\ln 2} \tag{5.5}$$

Given the required bandwidth, for both spatial and temporal frequency, σ_f can be obtained by numerically solving the equation.

An example of the receptive fields (in image space) of the V1 population of neurons used to simulate the responses of one particular MT neuron is illustrated in Figure 5.3. Since only the sine component of the quadrature pair is plotted for each combination of optimal spatial and temporal frequency, the figure shows 7x7=49 receptive fields (half of the total V1 filter population).

The Fourier representations of the receptive fields shown in Figure 5.3 are plotted in Figure 5.4 as a series of spatial frequency tuning curves, and Figure 5.5 as a series of temporal frequency tuning curves. Since V1 receptive fields in the model are spatiotemporally separable, these tuning curves completely characterise the receptive fields in the Fourier space.

The spatial frequency tuning curves (each curve corresponds to the spatial frequency tuning of one column of neurons in Fig. 5.3) in Figure 5.4 are plotted in two formats: in linear scale (Fig 5.4a) and in logarithmic scale (Fig. 5.4b). The lower panel illustrates that although the optimal spatial frequencies are evenly distributed in logarithmic scale, they grow exponentially in linear scale. In addition, it can also be seen that although the spatial

frequency tuning bandwidth, measured in octave (a logarithmic unit), decreases with the optimal spatial frequency, in linear scale, the range of spatial frequency that the neuron

responds to actually increases with the optimal spatial frequency. The representation of temporal frequency selectivity is illustrated in the same format in Figure 5.5.

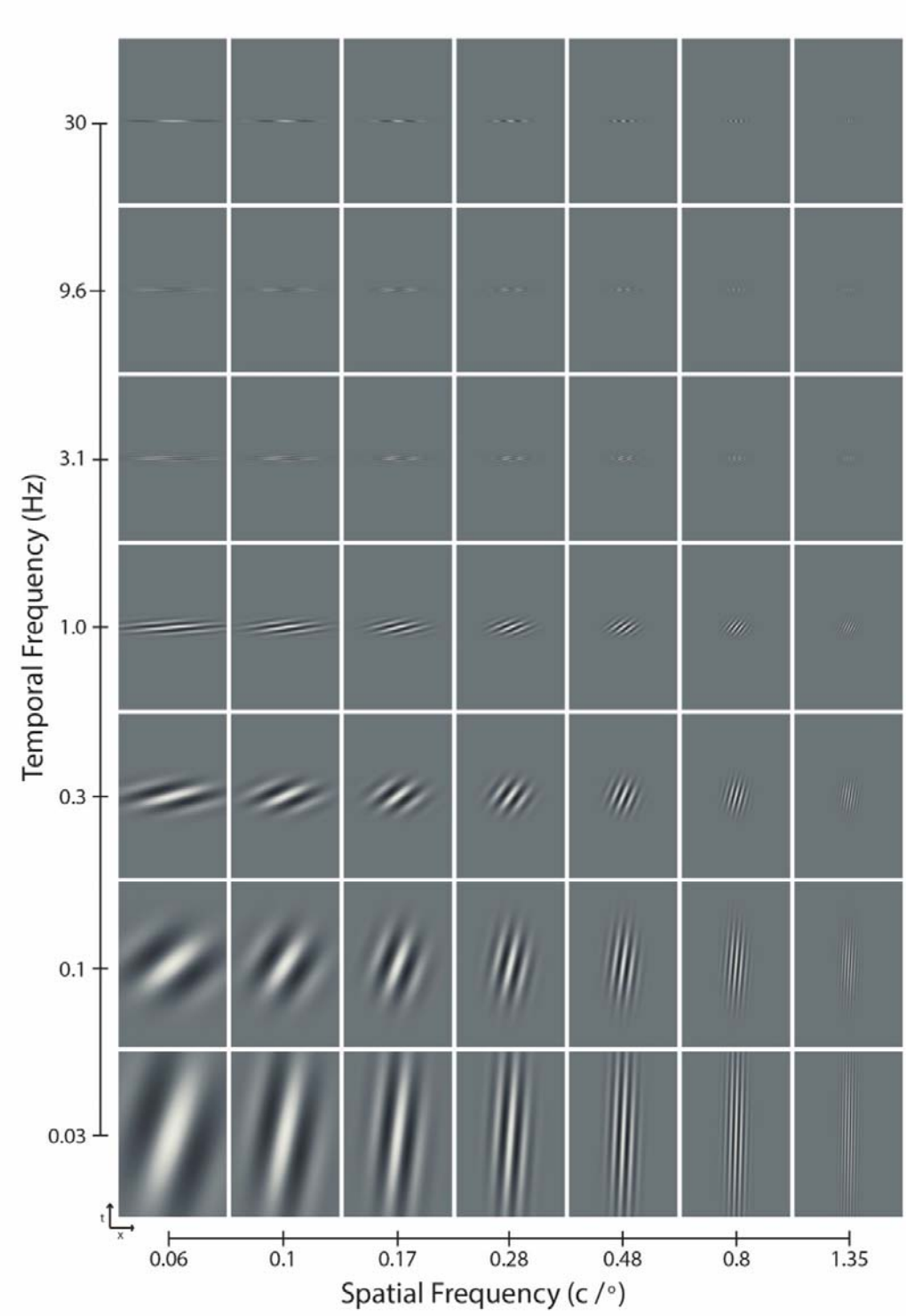

Figure 5.3. An example of the receptive fields of a modelled population of V1 neurons (showing only the sine component of the quadrature pair in the image space).

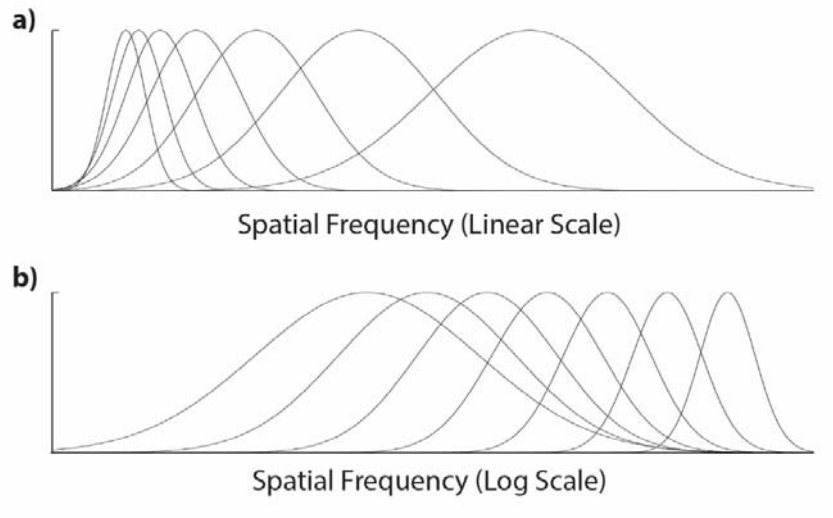

Figure 5.4. Spatial frequency tuning curves of the model V1 neurons shown in Figure 5.3. Each curve represents the spatial frequency selectivity of one column of neurons in Figure 5.3. The curves are represented in linear scale a) and in logarithmic scale b).

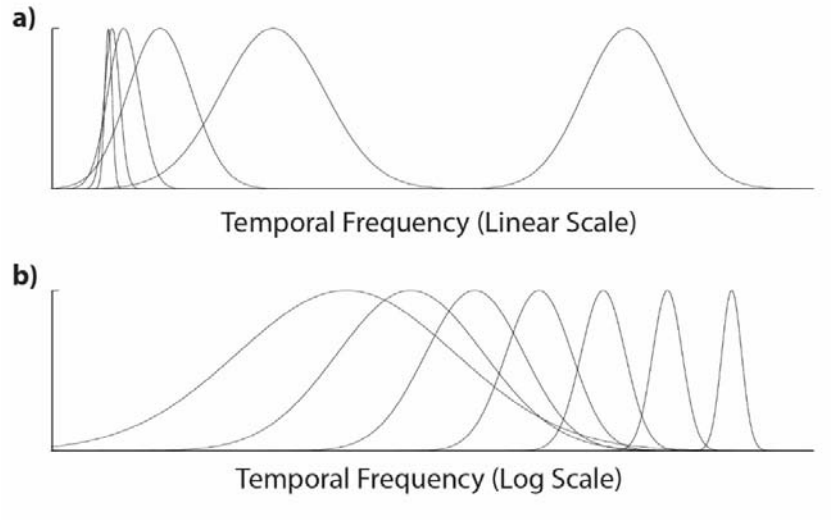

Figure 5.5. Temporal frequency tuning curves of the model V1 neurons shown in Figure 5.3. Each curve represents the temporal frequency selectivity of one row of neurons in Figure 5.3. The curves are represented in linear scale a) and in logarithmic scale b).

5.3.3 Nonlinear operations

To implement V1 neurons as "elementary motion detectors", the outputs of quadrature pairs of Gabor filters were squared and summed to calculate motion energy (Section 1.6.1). The nonlinear squaring operation removes the phase dependency of linear filters, resulting in the response characteristics of V1 "complex cells".

Since the stimulus always moved in the preferred direction of the neuron (never in the null direction), the motion opponency operation described in Section 1.6.1 was unnecessary and therefore not implemented.

Following the general architecture of the Simoncelli & Heeger (1998) model for MT (see also Nishimoto & Gallant, 2011), the output of the elementary motion detectors were normalised across the population (Section 1.6.2). This is the second nonlinear operation used in the model.

5.3.4 Modelling MT neurons by pooling V1 responses

The response pattern of a MT neuron is modelled as the linear sum of the (normalised) outputs of V1 neurons. The weights of the 49 V1 inputs were estimated by regression to approximate the tuning surface obtained with the Standard mode stimulus.

The responses to the Moving mode stimulus were then provided by the model. As in previous chapters, speed-tuning under Moving mode stimulus was determined by fitting the tuning surface to two models (M_0 and M_1; Chapter 2) and the model that was more likely to explain the data was selected with the Bayesian Information Criterion (BIC).

5.3.5 The stimuli

The Standard mode stimulus is essentially the spatiotemporal Gabor function described in Section 5.3.1, but without the temporal Gaussian envelope. It is mathematically described as

$$F_s(x,t) = e^{\left(-\pi\left(\frac{2}{rf}\right)*x\right)^2} * e^{i2\pi\,(sf0*x+tf0*t)} \tag{5.6}$$

where rf determines the size of the stimulus, sf_0 the central spatial frequency and tf_0 the central temporal frequency (since in the experiments, speed instead of temporal frequency was manipulated, tf_0 is converted from speed using $tf = speed * sf$). The first exponential function in the equation represents a Gaussian envelope, to avoid the sharp edges that are presented in many experiments based on sinusoidal grating patches. The second exponential represents the drifting sinusoidal carrier. Only the real part of the complex function $F_s(x,t)$ was used for simulating the stimulus, which is visualised as a two-dimensional plot in Figure 5.6, indicating that the Gabor pattern drifted in time (as spatial phase modulation), while the envelope remained stationary. To facilitate comparison to the Moving mode stimulus, in Figure 5.6 and 5.7, all the subplots in spatial coordinate have the same range. The range of x was set to $\{-rf, rf\}$, which represents the distance covered by the Moving mode stimulus. The range of t was limited to $\{-\frac{2rf}{sp_{min}}, \frac{2rf}{sp_{max}}\}$, (where sp_{min} and sp_{max} are the minimum and maximum values of speed selected for the population as outlined in Section 5.3.2) representing the longest possible time interval for the stimulus to pass completely through the receptive field in the Moving mode. In the simulation, the space domain was uniformly sampled at 200 points, and the time domain was sampled at 800 points.

The Moving mode stimulus is mathematically described as

$$F_m(x,t) = e^{(-\pi\left(\frac{2}{rf}\right)*(x+sp0*t))^2} * e^{i2\pi*sf0*(x+sp0*t)} \quad (5.7)$$

where rf represents stimulus size, sf_0 the central spatial frequency, and sp_0 the speed of the Gabor pattern. This stimulus is illustrated in Figure 5.7, which shows that the Gabor pattern moves in space, resulting in a slanted appearance in the x-t space.

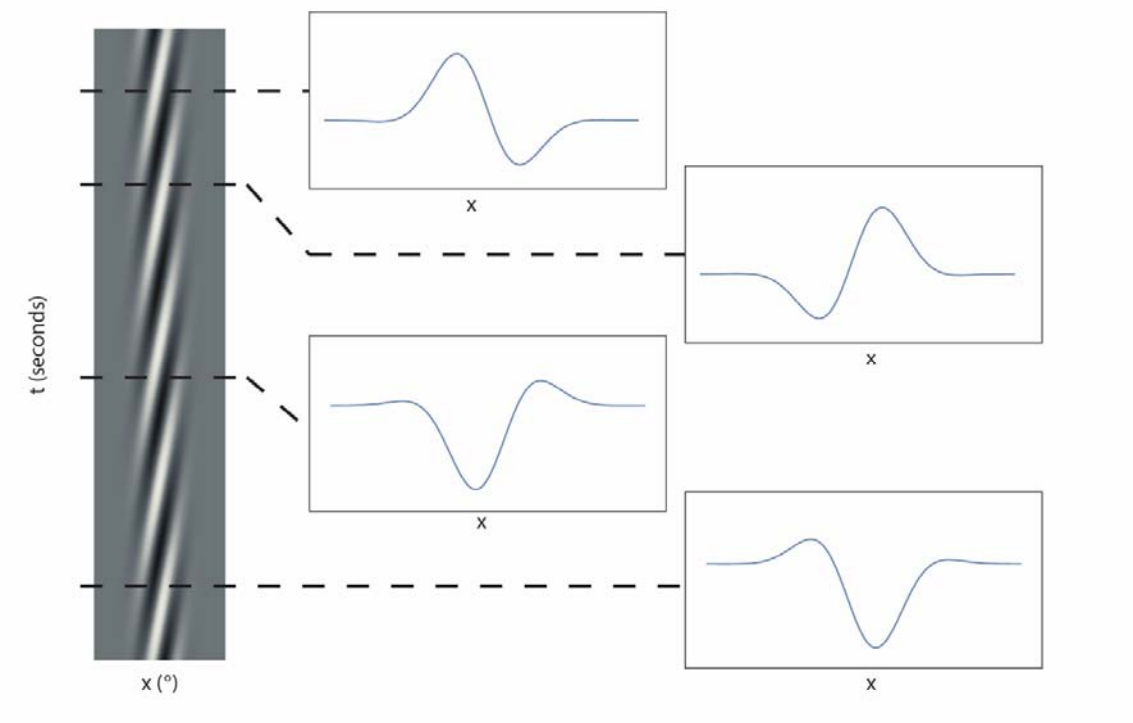

Figure 5.6. Visual representation of the Standard mode stimulus in the 2D space (left). The x-axis represents space. The y-axis represents time. When samples are taken at different points in time (inset) it can be appreciated that the carrier wave is drifting over time while the envelope remains stationary. This results in the diagonal-stripe appearance of the stimulus.

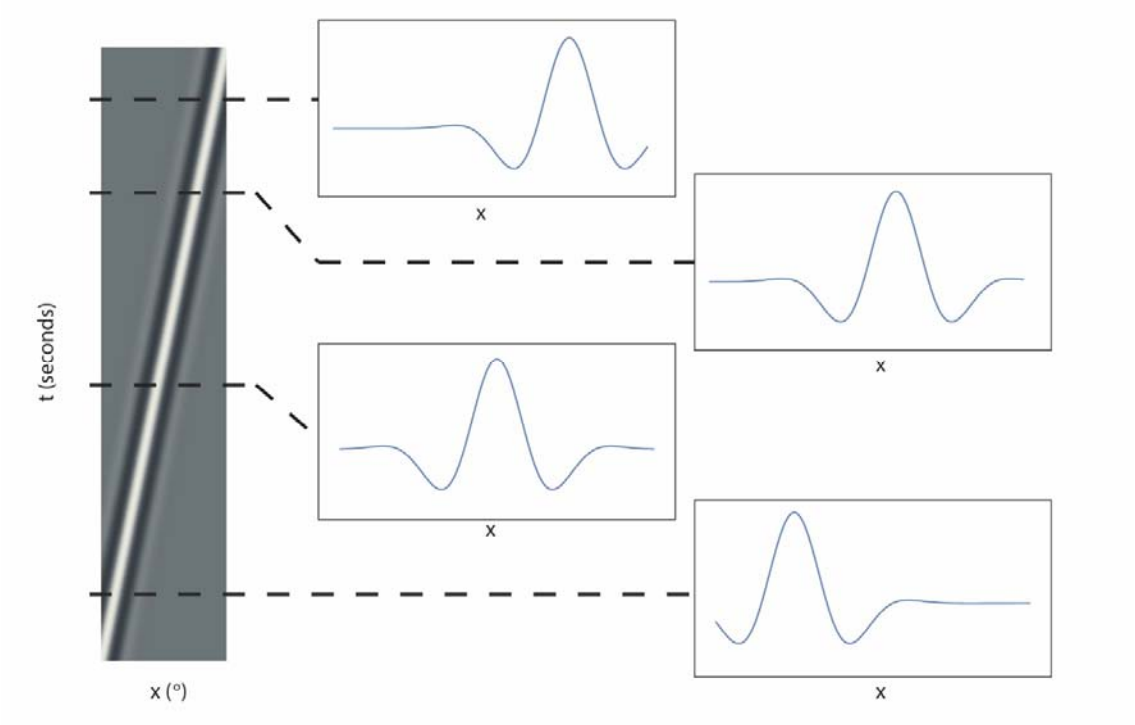

Figure 5.7. Visual representation of the Moving mode stimulus in 2D space (left). When samples are taken at different points in time (inset) it is possible to see that the Gabor pattern moves in space, resulting in the slanted appearance in the x-t plot.

5.4 Results

To examine if the results reported in Chapter 3 could be accounted for by a commonly accepted linear summation model for MT computation, we estimated the model parameters (i.e. the summation weights) for each of the 110 MT neurons using the tuning surfaces measured with the Standard mode stimulus, and then calculated the expected tuning surfaces when tested with the Moving mode stimulus.

5.4.1 Speed-tuning classifications of model neurons

Table 5.1 summarises the speed-tuning classifications of the modelled MT neurons. The model was able to correctly duplicate the speed-tuning classifications of all the 110 recorded MT neurons (second column of Table 5.1), measured in Standard mode. When the models were used to predict the speed-tuning classifications under the Moving mode, the classification of only one modelled neuron was changed from being non-speed-tuned to speed-tuned (fourth column of Table 5.1). The model was therefore unable to duplicate the main findings of Chapter 3, where 51 of the recorded MT neurons (46.4%) changed their classification from being non-speed tuned to speed-tuned when tested with the Moving mode stimulus (third column of Table 5.1).

Table 5.1. Speed-tuning classification of neurons across the two recorded, and two modelled conditions. Speed-tuning was determined using a threshold criterion of $P(M_0|M_0,M_1) > 0.5$. As there are 16 possible combinations of response patterns, for ease of viewing only categories that were found in this study are displayed. For all other categories, n = 0.

Recorded Standard Mode	Modelled Standard Mode	Recorded Moving Mode	Modelled Moving Mode	Number
Speed-Tuned	Speed-Tuned	Speed-Tuned	Speed-Tuned	1 (0.9%)
Non Speed-Tuned	Non Speed-Tuned	Non Speed-Tuned	Speed-Tuned	1 (0.9%)
Non Speed-Tuned	Non Speed-Tuned	Speed-Tuned	Non Speed-Tuned	51 (46.4%)
Non Speed-Tuned	Non Speed-Tuned	Non Speed-Tuned	Non Speed-Tuned	57 (51.8%)

Figure 5.8 to Figure 5.11 illustrate the tuning surfaces produced by the models: Figure 5.8 illustrates a case where the recorded neuron was speed-tuned in both the Standard mode and the Moving mode (First row of Table 5.1). The behaviour was correctly duplicated by the model. Figure 5.9 illustrates the only case where the recorded neuron was not speed-tuned in the Standard and in the Moving mode (second row of Table 5.1), but the model predicted speed-tuning in the Moving mode. Figure 5.10 and Figure 5.11 illustrate two representative cases that account for the majority of the sampled neuron: Figure 5.10 illustrates a representative case of neurons that were speed-tuned only under Moving mode (third row of Table 5.1) – a condition that was not duplicated by the model, and Figure 5.11 illustrates a representative case of neurons that were not speed-tuned under both modes – a condition that was duplicated by the model.

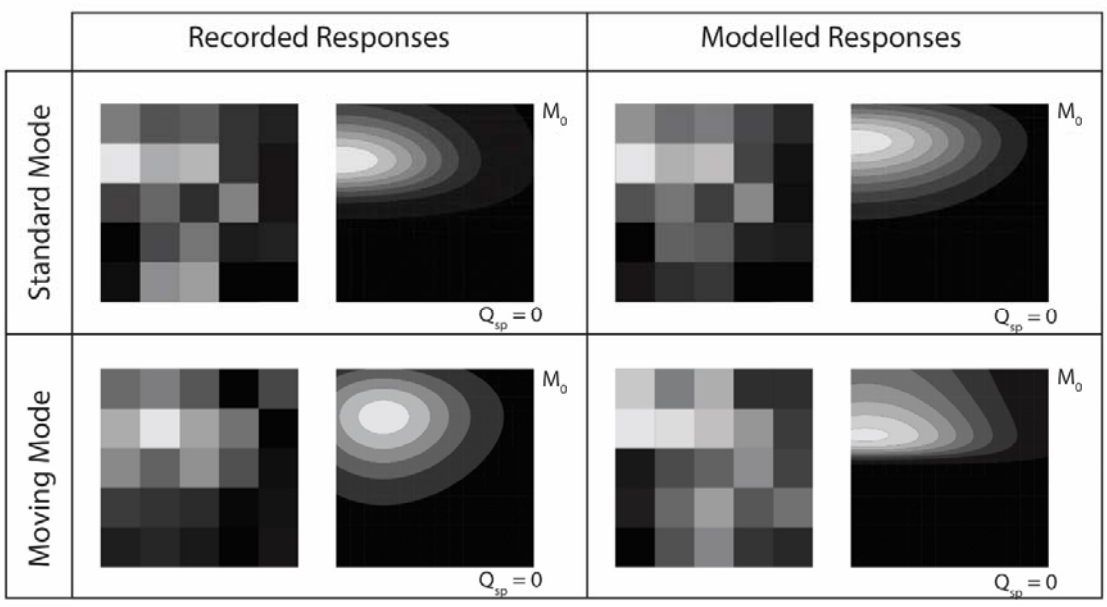

Figure 5.8. This figure shows a neuron that was found to be speed-tuned in all four conditions: both recorded modes, as well as both modelled responses (first row, Table 5.1). Within each condition, the matrix of responses is shown on the left, while the model selected as the better fit by $P(M_0|M_0,M_1)$ is shown on the right (model indicated on the top right corner). Q_{sp} value has also been provided for reference.

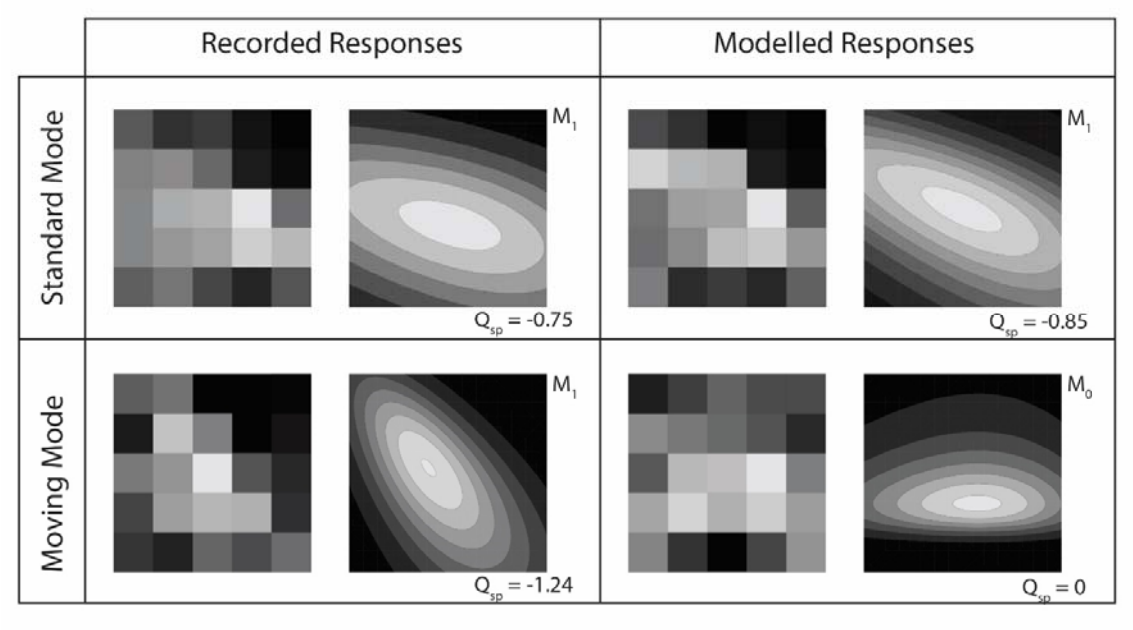

Figure 5.9. This figure shows a neuron that was found to be speed-tuned only under the modelled response to the Moving mode (second row, Table 5.1). The data are presented in the same format as Figure 5.8.

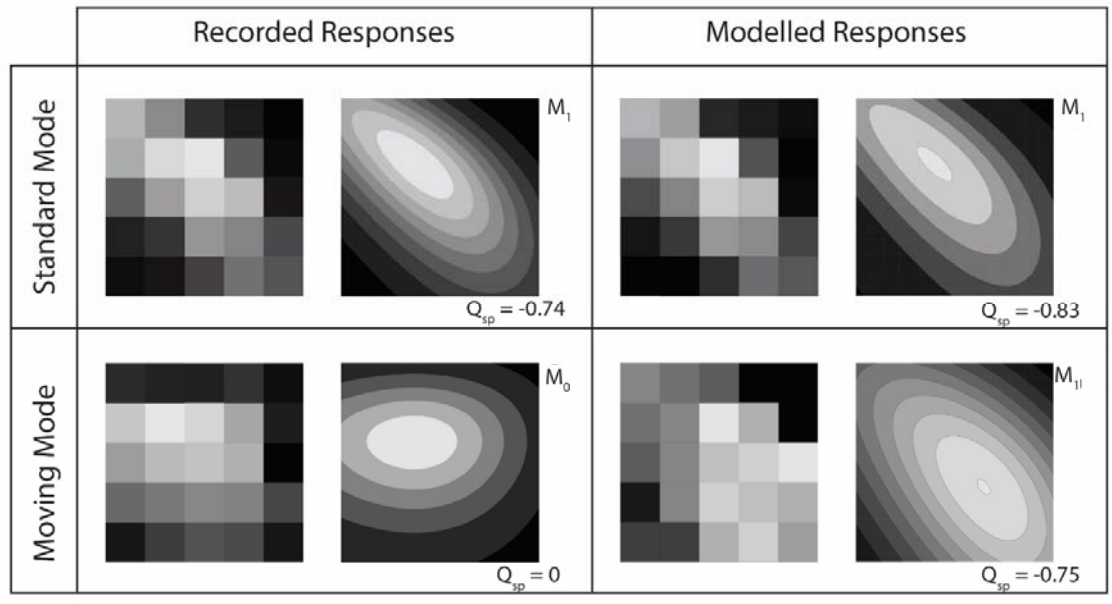

Figure 5.10. This figure shows a neuron that was found to be speed-tuned only under the recorded response to the Moving mode (third row, Table 5.1). The data are presented in the same format as Figure 5.8.

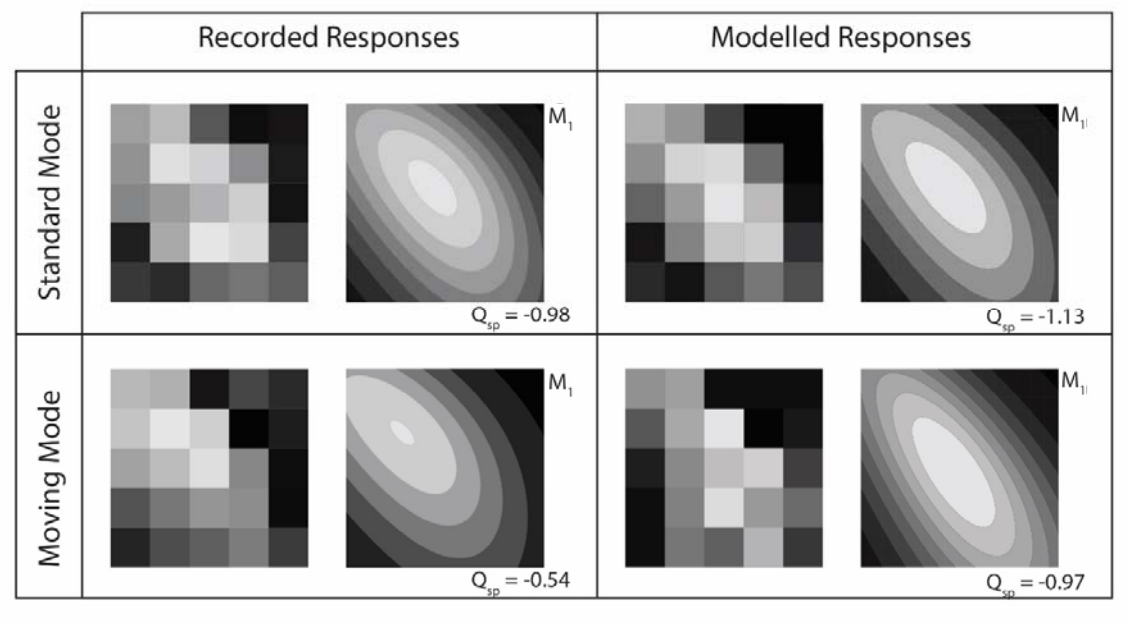

Figure 5.11. This figure shows a neuron that was found to be non speed-tuned in all four conditions: both recorded modes, as well as both modelled responses (fourth row, Table 5.1). The data are presented in the same format as Figure 5.8.

The classifications tabulated in Table 5.1 are also reflected in the distribution of $P(M_0|M_0,M_1)$, as shown in the right column of Figure 5.12. In the physiological recordings, $P(M_0|M_0,M_1)$ is bimodally distributed in the Moving mode (top row), indicating a significant population of speed-tuned neurons. However, the modelled response to the Moving mode produced a unimodal distribution clustered around 0 (middle row), which is more similar to the distribution of the experimental data under Standard mode (bottom row) than to the experimental data under Moving mode (top row).

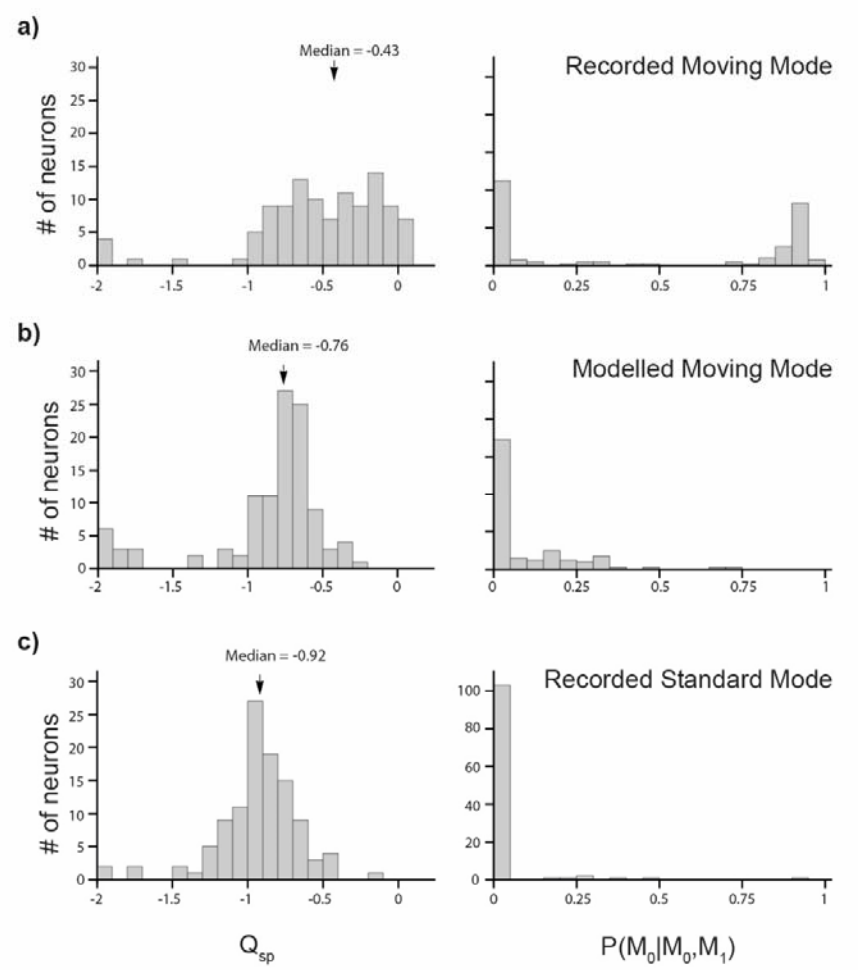

Figure 5.12. Population graphs showing the distribution of Q_{sp} and $P(M_0|M_0,M_1)$. Black arrows on the Q_{sp} graphs indicate population median values.

5.3.2 The response characteristics of individual model neurons

The median value of Q_{sp} for the recorded neuron under Standard mode was -0.92 (Fig. 5.12c, left panel). This value was increased to -0.76 when the constructed models were used to predict the responses to the Moving mode stimulus (Fig. 5.12b, left panel; Wilcoxon test: V = 2132, P = 0.0061), indicating that the linear pooling model predicted a small but significant shift towards speed-tuning under the Moving mode stimulus. However, the shift was not enough to account for the distribution obtained using the experimental data (Fig. 5.12a, left panel. Median value = -0.43; Wilcoxon test: V = 5079, P<0.0001).

The distributions of Q_{sp} are also visualised on an individual neuron basis in Figure 5.13. Figure 5.13a illustrates the model predicted (x-axis) and experimental (y-axis) Q_{sp} values for the Moving mode, where it is possible to see that a large part of the population of neurons lie above the dotted line (n = 85), indicating the models systematically underestimated SF-speed dependency for most of the neurons. Neurons indicated by white dots were found to be speed-tuned in the experimental data, but were not speed-tuned in the predictive model (n = 51). The figure highlights two sub-populations of neurons: those that became speed-tuned in the experimental data, and those which did not change classification. Of the second sub-population the model does a fairly accurate job of predicting the response to the Moving mode, with many neurons lying on, or close to, the dashed line, indicating identical responses for both conditions.

The upper panel of Figure 5.13b, plotting the model predicted Q_{sp} values in the Moving mode against those obtained from recorded data in the Standard mode, indicates that for the majority of the neurons (n = 72), the Moving mode stimulus increased the predicted values of Q_{sp} (i.e., increased the degree of speed-tuning). The median value of the increase was 0.13 (Figure 5.13b, lower panel). Although it was a moderate increase, the increments were not large enough to turn them into speed-tuned neurons (except for one single case). This is very different from the Q_{sp} values obtained from the recorded data (Figure 5.13c).

These results suggest that that while the model predicts small reductions in the dependency between speed and spatial frequency, the difference was not sufficient to explain the results presented in Chapter 3.

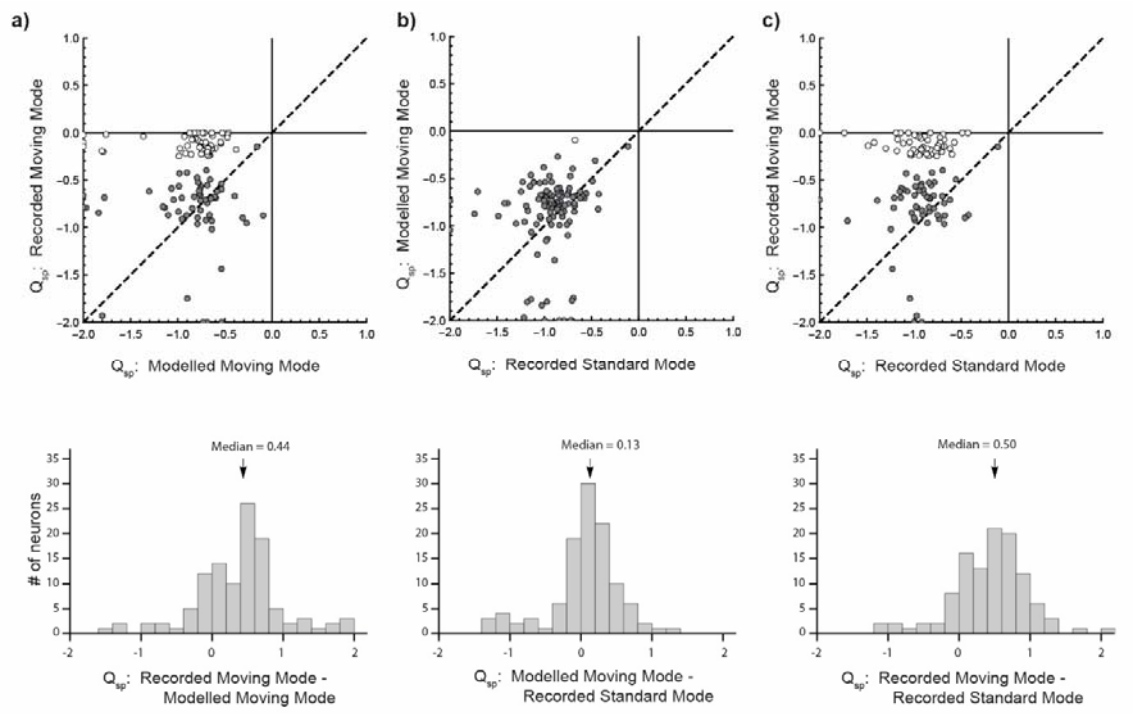

Figure 5.13. The distributions of Q_{sp} on an individual neuron basis. Upper panels: a) Comparison between the Q_{sp} in Moving mode, as predicted by the model, against the experimental values. b) Comparison between the predicted value for the Moving mode against the experimental values in the Standard mode. c) Comparison between the experimental values in the Standard mode against those in the Moving mode. White circles represent neurons that were changed from being non-speed-tuned (in the mode specified on on the x-axis) to being speed-tuned (in the mode specified on the y-axis). Dark grey circles represent all other neurons. The dashed lines represent values where Q_{sp} was identical in the two modes. Lower panels: The distributions of the difference of Q_{sp} between modes.

5.4 Discussion

By constructing computational models from data obtained with the Standard mode stimulus, we predicted the responses to the Moving mode stimulus, and compared the results against the recorded data. The model produced only 1 case (1%) where a non-speed tuned neuron under the Standard mode became speed-tuned under Moving mode. While this number has dramatically underestimated the increased number of speed-tuned neurons found within the Moving mode in the recorded data (51 neurons, 46%), it shows that such behaviour is theoretically possible.

As argued in Section 5.2, for a completely linear model, the responses to the Moving mode stimulus should produce tuning surfaces that are identical to those produced with the Standard mode stimulus. However, a comparison of the Q_{sp} parameter between the Standard mode and the Moving mode suggests that the three non-linear factors outlined in Introduction did have an effect on the speed-tuning behaviour in the Moving mode. Fig. 5.13b indicates that when stimulated with the Moving mode stimulus, the model neurons' Q_{sp} parameter was typically increased, meaning that there was a reduction in the dependency between speed-tuning and the spatial frequency contents of the stimulus. The magnitude of this reduction was 0.13 (Fig. 5.13b lower panel). While this reduction was a significant effect, it dramatically underestimated the shift in Q_{sp} in the recorded data (Fig. 13c), where a median magnitude of 0.5 was observed.

From these observations, we conclude that the stimulus-dependent speed-tuning results reported in Chapter 3 were unlikely to be explainable solely by the Fourier composition of the Moving mode stimulus. In other words, the responses to the Moving mode stimulus were different from those to the Standard mode stimulus, not simply because the stimuli were different. Rather, a form of nonlinear summation, which is not present in the model, was likely needed to explain the phenomenon.

Further studies are needed to determine the computational principles that might explain the results. Some promising directions are:

1. Bayesian models are probabilistic models, which estimate the most likely interpretation for a given set of sensory data. Importantly, the estimation of likelihood in Bayesian models is influenced by prior knowledge about the current state of the world. Bayesian models have been used in the past to explain how the coding of speed and direction can be altered by the properties of the stimulus (Ascher & Grzywacz, 2000; Weiss et al., 2002). The trajectory of the Moving mode stimulus might be able to alter the prior probabilities in a Bayesian model, which might then be able to account for the change in shape of the tuning surface.

2. The computational model implemented in this chapter is a form of feed-forward model, meaning that a MT neuron receives inputs from units in a lower hierarchal level (V1), but not from itself or from other MT neurons. Although feed-forward models have been used successfully in modelling the visual cortex, and will remain popular due to their analytical simplicity, these type of models are "memory-less", meaning that the passage of time is not explicitly represented in the model, and activities of the past are not permitted to influence the behaviour of the model at a later time point. This level of simplification is ill-suited for modelling the responses to the Moving mode stimulus, since the temporal history of the stimulus is an important aspect of the motion trajectory.

Recurrent models are computational models where the activity of a model MT neuron can feedback to itself or to other MT neurons, and therefore can influence their responses at a later time point. Recurrent models have been used to model MT neurons (e.g., Joukes et al., 2014 "motion detection based on recurrent network dynamics"), and might provide a basis for explaining the Moving mode data.

3. Motion trajectories contain important contextual information because they provide predictions of the location of moving objects in the future. Several psychophysical experiments have suggested that a neuron that detects local motion signals can send "forward signals" along the path of the trajectory (McKee & Welch, 1985; Grzywacz et al., 1995; Roach et al., 2001; Neri, 2014), which can be used to alter the behaviour of other neurons at a future time point. To implement this type of models, multiple MT neurons will need to communicate with each other according to the locations of their receptive fields. It is therefore a specific type of recurrent model where the connections are constrained by the visuotopy of the receptive fields.

This paper together with the paper of Priebe et al., (2003; 2006) has shown that the linear pooling approach has significant problems in estimating speed-tuning within area MT. So far there remains to be a model that robustly accounts for these nonlinear phenomena. In particular, my study involves a type of contextual effect introduced by including a stimulus trajectory. And while very little attempt has been made to date in order to account for these effects, these studies have shown that it is important to develop a model to account for these nonlinearities, particularly when investigating motion processing.

6. Discussion

The neural processing of moving images within the visual cortex is one of the most intensively studied topics in visual neuroscience. However, despite tremendous effort, the majority of the research has been focused on the encoding of the direction of motion. There is surprisingly little known about the mechanisms behind the processing of speed. In particular, much is unknown about how neurons in the primate visual cortex are tuned to the speed of moving objects in a way that is not dependent on the composition of the spatial pattern of the moving object. The Middle Temporal (MT) area has been considered one of the centres of motion processing in the primate visual cortex. Yet even within this well-studied area, the extent to which neurons encode speed remains a controversial issue.

In my thesis, I used a combination of electrophysiological recordings and computational modelling to address several issues that are central to this problem. An innovative feature of my study was that while previous studies probed the response characteristics of neurons with localised drifting patterns confined to the receptive fields of individual neurons, I used patterns that moved across the visual field in trajectories, which created temporal context of the motion signals, similar to how objects move in natural environments. My thesis is therefore also an approach for studying the effects of context in motion processing in the

visual cortex. While it has long been recognised that neuronal responses in the visual cortex are very often context-dependent, the effects of the temporal context of moving objects have not been examined at single neuron level.

The main results are described in Chapter 3, 4, and 5. In Chapter 3, I revisited the problem of speed-tuning in area MT of the marmoset monkey. Area MT is often considered the centre of motion processing within primate visual cortex. So when electrophysiological studies began to look for neurons capable of signalling motion speed, regardless of the spatial frequency composition of the stimulus, area MT is an area that was frequently targeted (Priebe et al., 2003; Lui et al., 2007; Miura et al., 2014). However, despite the ecological importance of accurate speed perception for survival (Gibson, 1950), most previous studies have reported small populations of speed-tuned neurons in area MT. In the marmoset in particular, as few as 10% of speed-tuned MT neurons have been reported (Lui et al., 2007). The puzzling findings suggest that the need for a large population of speed-tuned neurons might been have overestimated. Perhaps temporal-frequency tuning is sufficient for supporting most of the visual functions. Or alternatively, it is also possible that speed is extracted at a later processing stage in the visual cortex (such as area MST).

In addition to using the drifting sinusoidal gratings that have been commonly used in previous studies, I quantified the speed-tuning characteristics of MT neurons using patterns that moved in trajectories (the Moving mode stimulus). Almost half of the neurons that were tested with both a drifting and moving stimulus were found to be speed-tuned only under the moving stimulus. The results demonstrated that speed-tuning in area MT is stimulus-dependent for a large population of neurons. The temporal context of a moving object in particular, has a significant effect on how motion is processed in MT. The results of previous studies that relied on drifting gratings must therefore be interpreted with caution, because the stimuli used likely did not reveal the full capacity of speed-processing in MT. My results are

consistent with that of Perrone & Thiele (2001) and Nishimoto & Gallant (2011), suggesting that as a centre of motion processing in the visual cortex, area MT contains large populations of neurons that encode both the speed and direction of moving patterns.

The visual cortex is commonly viewed as a hierarchical network. Under this view, the properties of a visual area at higher levels of the hierarchy are a combination of those inherited from a lower level, as well as unique properties that are created de novo by the local circuits within that area. As such, it is important to establish which properties are inherited, and which are unique. For area MT, the largest input originates in the primary visual area (V1), and these inputs are thought to be responsible for the ubiquitous presence of direction-tuned neurons within MT. This led to the research described in Chapter 4, where I used electrophysiological recordings to investigate the response of V1 neurons to the same stimuli that were used in MT. In Chapter 4, I reported a small number of neurons (n = 6) that were found to be speed-tuned under the moving condition. However, for the vast majority (95%) of neurons that were tested, this was not the case: they remained non-speed tuned when tested with the Moving mode stimulus. The result is in dramatic contrast to area MT, where about half of the neurons became speed-tuned under the same condition. The results suggest that unlike direction tuning in MT, which has been suggested to be derived from the V1, the contextual effect reported in Chapter 3 was not inherited from V1 neurons.

The most detailed previous study on the origin of speed-tuning in area MT has been Priebe et al. (2003), who characterised speed-tuning in the macaque V1 and MT, using traditional drifting gratings. They came to the conclusion that a moderate population (~25%) of MT neurons are speed-tuned, which are likely to inherit their speed-tuning from V1. My results are inconsistent with that conclusion. Instead, I argued that a large population of MT neurons are speed-tuned (~50%) under the right context (where motion trajectories are present), and that this property was not inherited from V1. My research therefore reinforces the concept of

hierarchical processing in the visual cortex, and calls into question the alternative view which argues that many of the characteristics of MT neurons are merely aggregations of characteristics that are already present in V1.

Was the contextual effect described in Chapter 3 a novel discovery, unexpected by current theory of MT computation; or was it the expected behaviour of current theories? In other words, did the results reveal anything new about MT neurons, or were the results merely due to the composition of the Moving mode stimulus, which has not been used previously to study speed-tuning? This question was addressed in Chapter 5 with computational modelling. A class of computational model, which linearly pools inputs from V1, has been successful in explaining a range of characteristics of MT neurons (Simoncelli & Heeger, 1998; Nishimoto & Gallant, 2011). In Chapter 5, I implemented such a model to predict the responses to the Moving mode stimulus, based on the measured responses to the drifting grating stimulus.

The model was found to allow for some degree of change in the spatiotemporal response profiles of MT neurons in the Moving mode. However, it grossly underestimated the amount of change found in the experimental results of Chapter 3, demonstrating only one neuron which was expected to change its speed-tuning classification from being non-speed tuned in the Standard mode to being speed-tuned in the Moving mode. The linear-pooling model was therefore unable to account for the responses of MT neurons to the Moving mode stimulus. The result calls for a re-evaluation of the popular linear-pooling model, and suggests a direction which incorporate context-dependent non-linear summation elements. The result is consistent with Priebe (2003), who also reported deviations from the predictions of linear summation models, except that in Priebe (2003), the non-linear summation effect investigated was in the spatial domain, whereas in my study, a form of nonlinear temporal summation was demonstrated.

This thesis has investigated the effect of a stimulus trajectory on the response properties of neurons, showing that the presence of temporal and spatial context can alter how many neurons within the area MT will respond to stimulus speed. This has highlighted the importance of considering how the context of the stimulus may affect motion processing, by showing that the processing of speed within the brain involves more complicated mechanisms than previously thought.

Reference List

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. *JOSA A*, *2*, 284-299.

Adelson EH, Bergen JR (1986, May). The extraction of spatiotemporal energy in human and machine vision. In *Proc. IEEE Workshop on Visual Motion* (pp. 151-156).

Ahmed B, Cordery PM, McLelland D, Bair W, Krug K (2012) Long-range clustered connections within extrastriate visual area V5/MT of the rhesus macaque. *Cereb Cortex*, *22*, 60-73.

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol 52,* 1106 –1130.

Albright TD, Desimone R (1987) Local precision of visuotopic organisation in the middle temporal area (MT) of the macaque. *Exp Brain Res,* 65, 582-592.

Albright TD, Desimone R, Gross CG (1984) Columnar organization of directionally selective cells in visual area MT of the macaque. *J Neurophys*, *51*, 16-31.

Allman JM, Kaas JH (1971) A representation of the visual field in the caudal third of the middle temporal gyrus of the owl monkey (Aotus trivirgatus). *Brain Res*, *31*, 85-105.

Allman JM, Miezin F, McGuinness E (1985) Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, *14*(2), 105-126.

Anderson JC, Binzegger T, Martin KA, Rockland KS (1998) The connection from cortical area V1 to V5: a light and electron microscopic study. *J Neurosci*, *18*(24), 10525-10540.

Anderson JC, Martin KA (2002) Connection from cortical area V2 to MT in macaque monkey. *J Comp Neurol*, *443*(1), 56-70.

Ascher D, Grzywacz NM (2000) A Bayesian model for the measurement of visual velocity. *Vision research*, *40*(24), 3427-3434.

Baizer JS, Ungerleider LG, Desimone R (1991) Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *J Neuro*, *11*, 168-190.

Baker CL, Jr. (1990) Spatial- and temporal-frequency selectivity as a basis for velocity preference in cat striate cortex neurons. *Vis Neurosci, 4,* 101-113.

Barlow HB (1961) The coding of sensory messages. *Current Problems in Animal Behaviour*, 331-360.

Born RT, Bradley DC (2005) Structure and function of visual area MT. *Annu. Rev. Neurosci.*, *28*, 157-189.

Bourne JA (2010) Unravelling the development of the visual cortex: implications for plasticity and repair. *J Anatomy*, *217*, 449-468.

Bourne JA, Rosa MG (2003) Preparation for the in vivo recording of neuronal responses in the visual cortex of anaesthetised marmosets (Callithrix jacchus). *Brain Research Protocols, 11,* 168-177.

Bourne JA, Tweedale R, Rosa MGP (2002) Physiological responses of new world monkey V1 neurons to stimuli defined by coherent motion. *Cereb Cortex, 12*, 1132-1145.

Bradley DC, Goyal MS (2008) Velocity computation in the primate visual system. *Nat Rev Neurosci*, *9*(9), 686-695.

Bremmer F, Schlack A, Duhamel JR, Graf W, Fink GR (2001) Space coding in primate posterior parietal cortex. *Neuroimage*, *14*(1), S46-S51.

Britten KH, Shadlen MN, Newsome WT, Movshon JA (1993) Responses of neurons in macaque MT to stochastic motion signals. *Vis Neurosci*, *10*(6), 1157-1169.

Burr D, Thompson P (2011) Motion psychophysics: 1985–2010. *Vision Res*, *51,* 1431–1456.

Carandini M, Heeger DJ (2012) Normalization as a canonical neural computation. *Nat Rev Neurosci*, *13*(1), 51-62.

Carandini M, Heeger DJ, Movshon JA (1997) Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci*, *17*(21), 8621-8644

Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol, 88,* 2530-2546.

Chaplin TA, Yu HH, Soares JG, Gattass R, Rosa MGP (2013) A conserved pattern of differential expansion of cortical areas in simian primates. *J Neuro*, *33*(38), 15120-15125.

Cheng K, Hasegawa T, Saleem KS, Tanaka K (1994) Comparison of neuronal selectivity for stimulus pseed, length and contrast in prestriate visual cortical areas V4 and MT of the macaque monkey. *J Neurophysiol, 71*(6), 2269-2280.

Clarke HF, Dalley JW, Crofts HS, Robbins TW, Roberts AC (2004) Cognitive inflexibility after prefrontal serotonin depletion. *Science, 304,* 878–880

Clarke HF, Hill GJ, Robbins TW, Roberts AC (2011) Dopamine, but not serotonin, regulates reversal learning in the marmoset caudate nucleus. *J Neurosci*, *31,* 4290–4297

Colby CL, Duhamel JR, Goldberg ME (1993) Ventral intraparietal area of the macaque: anatomic location and visual response properties..*J Neurophys*, *69*(3), 902-914.

Daugman JG (1985) Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *JOSA A*, *2*(7), 1160-1169.

Davies AJ, Chaplin TA, Rosa MGP, Yu HH (2016) Natural motion trajectory enhances the coding of speed in primate extrastriate cortex. *Scientific Reports*, *6*.

de Haan E H, Cowey A (2011) On the usefulness of 'what'and 'where'pathways in vision. *Trends in Cognitive Sciences*, *15*(10), 460-466.

De Valois RL, Albrecht DG, Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*(5), pp.545-559.

De Valois RL, Yund EW, Hepler N (1982) The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, *22*(5), 531-544.

Desimone R, Albright TD, Gross CG, Bruce C (1984) Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci*, *4*(8), 2051-2062.

Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophys*, *57*(3), 835-868.

Diogo ACM, Soares JG, Koulakov A, Albright TD, Gattass R (2003) Electrophysiological imaging of functional architecture in the cortical middle temporal visual area of Cebus apella monkey. *J Neurosci*, *23*, 3881-3898.

Dubner R, Zeki SM (1971) Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research*, *35*(2), 528-532.

Duffy CJ, Wurtz RH (1991) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J Neurophys*, *65*(6), 1329-1345.

Dukelow SP, DeSouza JF, Culham JC, van den Berg AV, Menon RS, Vilis T (2001) Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *J Neurophys*, *86*(4), 1991-2000.

Dumoulin SO, Bittar RG, Kabani NJ, Baker CL, Le Goualher G, Pike GB, Evans AC (2000). A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex*, *10*(5), 454-463.

Emerson RC, Bergen JR, Adelson EH (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, *32*(2), 203-218.

Felleman DJ, Kaas JH (1984) Receptive-field properties of neurons in the middle temporal visual area (MT) of owl monkeys. *J Neurophys 52*, 488–513.

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb cortex*, *1*, 1-47.

Ferreira CT, Ceccaldi M, Giusiano B, Poncet M (1998) Separate visual pathways for perception of actions and objects: evidence from a case of apperceptive agnosia. *Journal of Neurology Neurosurgery and Psychiatry, 65,* 382-385.

Fiorani M, Gattass R, Rosa MG, Sousa AP (1989) Visual area MT in the Cebus monkey: location, visuotopic organization, and variability. *J Comp Neuro*, *287*, 98-118.

Forte JD, Hashemi-nezhad M, Dobbie WJ, Dreher B, Martin PR (2005) Spatial coding and response redundancy in parallel visual pathways of the marmoset Callithrix jacchus. *Vis Neurosci*, *22*(04), 479-491.

Foster KH, Gaska JP, Nagler M, Pollen DA (1985) Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *J Physiol*, *365*(1), 331-363.

Fritsches KA, Rosa MGP (1996) Visutopic organisation of striate cortex in the marmoset monkey (Callithrix Jacchus). *J Comp Neurology*, *372*(2) 264-282.

Gallant JL, Braun J, Van Essen DC (1993) Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science*,*259*(5091), 100-103.

Gallyas F (1978) Silver staining of myelin by means of physical development. *Neurological Research*, *1*(2), 203-209.

Gattass R, Gross CG (1981) Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. *J Neurophys*, *46*(3), 621-638.

Gibson JJ (1950) The perception of the visual world.

Girard P, Bullier K (1989) Visual activity in area V2 during reversible inactivation of area 17 in the macaque monkey. *J Neurophys, 62*, 1287-1302.

Girard P, Salin PA, Bullier J (1991) Visual activity in macaque area V4 depends on area 17 input. *Neuroreport, 2*, 81-84.

Girard P, Salin PA, Bullier J (1992) Response selectivity of neurons in area MT of the macaque monkey during reversible inactivation of area V1. *J Neurophys, 67*, 1437-1446.

Goldberg ME, Bisley JW, Powell KD, Gottlieb J (2006) Saccades, salience and attention: the role of the lateral intraparietal area in visual behavior. *Progress in Brain Research*, *155*, 157-175.

Goodale MA (2001) Different spaces and different times for perception and action. *Progress in Brain Research*, *134*, 313-331.

Goodale MA (2011) Transforming vision into action. *Vision Research*, *51*, 1567-1587.

Goodale MA, Króliczak G, Westwood DA (2005) Dual routes to action: contributions of the dorsal and ventral streams to adaptive behavior. *Progress in Brain Research*, *149*, 269-283.

Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. *Trends in Neuroscience, 15*, 20-25.

Goodale MA, Westwood DA (2004) An evolving view of duplex vision: separate but interacting cortical pathways for perception and action. *Current Opinion in Neurobiology, 14*, 203-211.

Grzywacz NM, Watamaniuk SN, Mckee SP (1995) Temporal coherence theory for the detection and measurement of visual motion. *Vision Research*, *35*(22), 3183-3203.

Guo K, Benson PJ, Blakemore C (2004) Pattern motion is present in V1 of awake but not anaesthetized monkeys. *Eur J Neurosci*, *19*(4), 1055-1066.

Gur M, Kagan I, Snodderly DM (2005) Orientation and direction selectivity of neurons in V1 of alert monkeys: functional relationships and laminar distributions. *Cereb Cortex*, *15*(8), 1207-1221.

Hearn JP (1983) "The common marmoset (Callithrix jacchus)." *Reproduction in New World Primates*. Springer Netherlands, 181-215.

Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. *J Physiol*, *148*(3), 574-591.

Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol*, *195*(1), 215-243.

Hung CC, Yen CC, Ciuchta, JL, Papoti D, Bock NA, Leopold DA, Silva AC (2015) Functional MRI of visual responses in the awake behaving marmoset. *Neuroimage, 120*, 1-11.

Ilg UJ (2008) The role of areas MT and MST in coding of visual motion underlying the execution of smooth pursuit. *Vision Research*, *48*(20), 2062-2069.

Jones JP, Palmer LA (1987) An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophys*, *58*(6), 1233-1258.

Joukes J, Hartmann TS, Krekelberg B (2014) Motion detection based on recurrent network dynamics. *Frontiers in Systems Neuroscience*, *8*, 239.

Kaas JH, Morel A (1993) Connections of visual areas of the upper temporal lobe of owl monkeys: the MT crescent and dorsal and ventral subdivisions of FST. *J Neurosci*, *13*(2), 534-546.

Kobatake E, Tanaka K (1994) Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophys*, *71*(3), 856-867.

Kravitz DJ, Saleem KS, Baker CI, Mishkin M (2011) A new neural framework for visuospatial processing. *Nat Rev Neurosci,,12*(4), 217-230.

Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M (2013) The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends in Cognitive Sciences*, *17*(1), 26-49.

Krubitzer LA, Kaas JH (1990) Cortical connections of MT in four species of primates: areal, modular, and retinotopic patterns. *Vis Neurosci*, *5*, 165-204.

Kulikowski JJ, Vidyasagar TR (1986) Space and spatial frequency: analysis and representation in the macaque striate cortex. *Experimental Brain Research*, *64*(1), pp.5-18.

Laan A, Gutnick T, Kuba MJ, Laurent G (2014) Behavioral analysis of cuttlefish traveling waves and its implications for neural control. *Current Biology*, *24*(15), 1737-1742.

Lagae S, Raiguel S, Orban GA (1993) Speed and direction selectivity of macaque middle temporal neurons. *J Neurophys, 69*, 19–39.

Lingnau A, Ashida H, Wall MB, Smith AT (2009) Speed encoding in human visual cortex revealed by fMRI adaptation. *J Vision*, *9*(13), 3-3.

Livingstone M, Hubel D (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*, 740-749.

Livingstone MS, Pack CC, Born RT (2001) Two-dimensional substructure of MT receptive fields. *Neuron*, *30*(3), 781-793.

Lui LL, Bourne JA, Rosa MGP (2007a). Spatial and temporal frequency selectivity of neurons in the middle temporal visual area of new world monkeys (Callithrix jacchus). *Eur J Neurosci*, *25*(6), 1780-1792.

Lui LL, Bourne JA, Rosa MGP (2007b) Spatial summation, end inhibition and side inhibition in the middle temporal visual area (MT). *J Neurophys, 97*(2), 1135-1148.

Lui LL, Rosa MGP (2015) Structure and function of the middle temporal visual area (MT) in the marmoset: comparisons with the macaque monkey. *Neurosci Res*, *93*, 62-71.

MacDougall M, Nummela SU, Coop S, Disney A, Mitchell JF, Miller CT (2016) Optogenetic manipulation of neural circuits in awake marmosets. *J Neurophys,* jn-00197.

Malach R, Schirman TD, Harel M, Tootell RB, Malonek D (1997) Organization of intrinsic connections in owl monkey area MT. *Cereb cortex*, *7*(4), 386-393.

Malonek D, Tootell RBH, Grinvald A (1994) Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT. *Proceedings of the Royal Society of London B: Biological Sciences*, *258*(1352), 109-119.

Maunsell JH, van Essen DC (1983a) The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neuro*, *3*(12), 2563-2586.

Maunsell JH, van Essen DC (1983b) Functional properties of neurons in middle temporal visual are of the macaque monkey. I. Selectivity for stimulus direction, speed and orientation. *J Neurophys*, *49*(5), 1127-1147.

Maunsell JH, van Essen DC (1987) Topographic organization of the middle temporal visual area in the macaque monkey: representational biases and the relationship to callosal connections and myeloarchitectonic boundaries. *J Comp Neur*, *266*, 535-555.

McKee SP, Silverman GH, Nakayama K (1986) Precise velocity discrimination despite random variations in temporal frequency and contrast. *Vision Research*, *26*(4), 609-619.

McKee SP, Welch L (1985) Sequential recruitment in the discrimination of velocity. *JOSA A*, *2*(2), 243-251.

Mikami A (1992) Spatiotemporal characteristics of direction-selective neurons in the middle temporal visual area of the macaque monkeys. *Experimental Brain Research*, *90*(1), 40-46.

Milner AD, Goodale MA (2008) Two visual systems re-viewed. *Neuropsychologia*, *46*, 774-785.

Mishkin M, Ungerleider LG, Macko KA (1983) Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, *6*, 414-417.

Mitchell JF, Leopold DA (2015) The marmoset monkey as a model for visual neuroscience. *Neuroscience Research, 93*, 20-46.

Mitchell JF, Reynolds JH, Miller CT (2014) Active vision in marmosets: a model system for visual neuroscience. *J Neurosci*, *34*(4), 1183-1194.

Miura K, Inaba N, Aoki Y, Kawano K (2014) Difference in visual motion representation between cortical areas MT and MST during ocular following responses. *J Neurosci 34*(6), 2160-2168.

Moore BD (2006) Speed selectivity in V1: A complex affair. *J Neurosci*, *26*, 7543-7544.

Movellan JR (2002) Tutorial on Gabor filters. *Open Source Document*.

Movshon JA, Adelson E, Gizzi M, Newsome W (1985) The analysis of moving visual patterns. In *Pattern Recognition Mechanisms*, (Pontificiae Academiae Scientiarum Scripta Varia) Vol. 54, 117–151 ed. Chagas C, Gattass R & Gross C. Vatican Press, Rome.

Movshon JA, Thompson ID, Tolhurst DJ (1978) Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex. *J Physiol*, *283*, p.101.

Movshon JA, Newsome WT (1996) Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J Neurosci*, *16*, 7733-7741.

Neri P (2014) Dynamic Engagement of Human Motion Detectors across Space–Time Coordinates. *J Neurosci.*, *34*(25), 8449-8461.

Newsome WT, Wurtz RH, Dursteler MR, Mikami A (1985) Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *J Neurosci*, *5*, 825-840.

Newsome WT, Pare EB (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *J Neurosci*, *8*, 2201-2211.

Ninomiya T, Sawamura H, Inoue KI, Takada M (2012) Segregated pathways carrying frontally derived top-down signals to visual areas MT and V4 in macaques. *J Neurosci*, *32*(20), 6851-6858.

Nishimoto S, Gallant JL (2011) A three-dimensional spatiotemporal receptive field model explains responses of area MT neurons to naturalistic movies. *J Neurosci*, *31*(41), 14551-14564.

Orban GA, Kennedy H, Maes H (1981) Response to movement of neurons in areas 17 and 18 of the cat: velocity sensitivity. *J Neurophys*, *45*(6), 1043-1058.

Orban GA, Kennedy H, Bullier J (1986) Velocity sensitivity and direction selectivity of neurons in areas V1 and V2 of the monkey: influence of eccentricity. *J Neurophys*, *56*(2), 462-480.

Palmer SM, Rosa MGP (2006) A distinct anatomical network of cortical areas for analysis of motion in far peripheral vision. *Eur J Neurosci*, *24*, 2389-2405.

Pasupathy A, Connor CE (2002) Population coding of shape in area V4. *Nat Neurosci*, *5*(12), 1332-1338.

Paxinos G, Watson c, Petrides M, Rosa MGP, Tokuno H (2012) *The marmoset brain in stereotaxic coordinates.* Elselvier.

Perrone JA, Thiele A (2001) Speed skills: measuring the visual speed analyzing properties of primate MT neurons. *Nat Neurosci*, *4*(5), 526-532.

Perrone JA, Thiele A (2002) A model of speed tuning in MT neurons. *Vision Research*, *42*(8), 1035-1051.

Priebe NJ, Cassanello CR, Lisberger SG (2003) The neural representation of speed in macaque area MT/V5. *J Neurosci*, *23*(13), 5650-5661.

Priebe NJ, Lisberger SG, Movshon JA (2006) Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *J Neurosci*, *26*(11), 2941-2950.

Reisbeck TE, Gegenfurtner KR (1999) Velocity tuned mechanisms in human motion processing. *Vision Research*, *39*(19), 3267-3286.

Roach NW, McGraw PV, Johnston A (2011) Visual motion induces a forward prediction of spatial pattern. *Current Biology*, *21*(9), 740-745.

Rockland KS (1989) Bistratified distribution of terminal arbors of individual axons projecting from area V1 to middle temporal area (MT) in the macaque monkey. *Visual Neuroscience*, *3*(02), 155-170.

Rockland KS (1995) Morphology of individual axons projecting from area V2 to MT in the macaque. *J Comp Neurol*, *355*(1), 15-26.

Rodman HR, Albright TD (1987) Coding of visual stimulus velocity in area MT of the macaque. *Vision Research, 27*, 2035–2048.

Rodman HR, Albright TD (1989) Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT).*Experimental Brain Research*, *75*(1), 53-64.

Rodman HR, Gross CG, Albright TD (1989) Afferent basis of visual response properties in area MT of the macaque. I. Effects of striate cortex removal. *J Neurosci*, *9*(6), 2033-2050.

Rosa MGP, Elston GN (1998) Visuotopic organisation and neuronal response selectivity for direction of motion in visual areas of the caudal temporal lobe of the marmoset monkey (*Callithrix jacchus*): middle temporal area, middle temporal crescent and surrounding cortex. *J Comp Neurol, 393*, 505-527.

Rosa MGP, Soares JG, Fiorani M, Gattass R (1993) Cortical afferents of visual area MT in the Cebus monkey: possible homologies between New and Old World monkeys. *Visual Neurosci*, *10*, 827-855.

Rosa MGP, Tweedale R (2005) Brain maps, great and small: lessons from comparative studies of primate visual cortical organisation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1456), 665-691.

Rosa MGP, Tweedale R, Elston GN (2000) Visual responses of neurons in the middle temporal area of new world monkeys after lesions of striate cortex. *J Neurosci*, *20*, 5552-5563.

Rust NC, Mante V, Simoncelli EP, Movshon JA (2006) How MT cells analyze the motion of visual patterns. *Nat Neurosci*, *9*(11), 1421-1431.

Rust NC, Schwartz O, Movshon JA, Simoncelli EP (2005) Spatiotemporal elements of macaque v1 receptive fields. *Neuron*, *46*(6), 945-956.

Roach NW, McGraw PV, Johnston A (2011) Visual motion induces a forward prediction of spatial pattern. *Current Biology*, *21*(9), 740-745.

Rodman HR, Gross CG, Albright TD (1989) Afferent basis of visual response properties in area MT of the macaque. I. Effects of striate cortex removal. *J Neurosci*, *9*, 2033-2050.

Rylands AB (1993) *Marmosets and tamarins: systematics, behavior, and ecology* (Oxford UP, Oxford, UK).

Sadakane O, Watakabe A, Ohtsuka M, Takaji M, Sasaki T, Kasai M, Isa T, Kato G, Nabekura K, Mizukami H, Ozawa K (2015) In vivo two-photon imaging of dendritic spines in marmoset neocortex, *eneuro, 2*(4), eneuro-0019.

Sasaki E, Suemizu H, Shimada A, Hanazawa K, Oiwa R, Kamioka M, Tomioka I, Sotomaru Y, Hirakawa R, Eto T, Shiozawa S (2009) Generation of transgenic non-human primates with germline transmission. *Nature*, *459*, 523-527.

Schiller PH, Malpeli JG (1977) The effect of striate cortex cooling on area 18 cells in the monkey. *Brain Research, 126*, 366-369.

Sengpiel F, Troilo D, Kind PC, Graham B, Blakemore C (1996). Functional architecture of area 17 in normal and monocularly deprived marmosets (Callithrix jacchus). *Visual Neuroscience*, *13*(01), 145-160.

Shapley R, Lennie P (1985) Spatial frequency analysis in the visual system. *Ann Rev Neurosci*, *8*(1), 547-581.

Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743-761.

Simoncelli EP, Heeger DJ (2001) Representing retinal image speed in visual cortex. *Nat Neurosci*, *4*(5), 461-462.

Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. *Ann Rev Neurosci*, *24*(1), 1193-1216.

Sincich LC, Park KF, Wohlgemuth MJ, Horton JC (2004) Bypassing V1: a direct geniculate input to area MT. *Nat Neurosci*,*7*(10), 1123-1128.

Smith AT, Edgar GK (1991) Perceived speed and direction of complex gratings and plaids. *JOSA A*, *8*(7), 1161-1171.

Smith MA, Majaj NJ, Movshon JA (2005) Dynamics of motion signaling by neurons in macaque area MT. *Nat Neurosci, 8*, 220–228.

Snowden RJ, Treue S, Anderson RA (1992) The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Exp Brain Res, 88*(2), 389-400.

Solomon SS, Tailby C, Gharaei S, Camp AJ, Bourne JA, Solomon SG (2011) Visual motion integration by neurons in the middle temporal area of a New World monkey, the marmoset. *J Physiol*, *589*(23), 5741-5758.

Solomon SG, Rosa MGP (2015) A simpler primate brain: the visual system of the marmoset monkey. *What can simple brains teach us about how vision works,* 136.

Spatz WB (1977) Topographically organized reciprocal connections between areas 17 and MT (visual area of superior temporal sulcus) in the marmoset Callithrix jacchus. *Exp Brain Res*, *27*(5), 559-572.

Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J Neurosci, 6,* 134-144.

Tinsley CJ, Webb BS, Barraclough NE, Vincent CJ, Parker A, Derrington AM (2003) The nature of V1 neural responses to 2D moving patterns depends on receptive-field structure in the marmoset monkey. *J Neurophys*, *90*(2), 930-937.

Tolhurst DJ, Thompson ID (1981) On the variety of spatial frequency selectivities shown by neurons in area 17 of the cat. *Proceedings of the Royal Society of London B: Biological Sciences*, *213*(1191), pp.183-199.

Touryan J, Lau B, Dan Y (2002) Isolation of relevant visual features from random stimuli for cortical complex cells. *J Neurosci*, *22*(24), 10811-10818.

Tsao DY, Vanduffel W, Sasaki Y, Fize D, Knutsen TA, Mandeville JB, Wald LL, Dale AM, Rosen BR, van Essen DC, Livingstone MS (2003) Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron*, *39*(3), 555-568.

Ungerleider LG, Desimone R (1986) Cortical connections of visual area MT in the macaque. *J Comp Neuro, 248*, 190-222.

Ungerleider LG, Mishkin M (1982) Two visual pathways (pp. 549–586).

Ungerleider LG, Haxby JV (1994) 'What' and 'where' in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157-165.

Van Essen DC, Maunsell JHR, Bixby JL (1981) The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *J Comp Neuro*, *199*(3), 293-326.

Watson AB, Ahumada AJ (1983) A look at motion in the frequency domain.

Watson AB, Ahumada AJ (1985) Model of human visual-motion sensing. *JOSA A*, *2*(2), 322-342.

Weiss Y, Simoncelli EP, Adelson EH (2002) Motion illusions as optimal percepts. *Nat Neurosci*, *5*(6), 598-604.

Wong-Riley M (1979) Changes in the visual system of monocularly sutured or enucleated cats demonstrable with cytochrome oxidade histochemistry. *Brain Research. 171*, 11-28.

Yamazaki Y, Echigo C, Saiki M, Inada M, Watanabe S, Iriki A (2011) Tool-use learning by common marmosets (Callithrix jacchus). *Exp Brain Res, 213*, 63–71.

Yu HH, Chaplin TA, Egan GW, Reser DH, Worthy KH, Rosa MGP (2013) Visually evoked responses in extrastriate area MT after lesions of striate cortex in early life. *J Neurosci*, *33*(30), 12479-12489.

Yu HH, Chaplin TA, Rosa MGP (2015) Representation of central and peripheral vision in the primate cerebral cortex: Insights from studies of the marmoset brain. *Neuroscience Research*, *93*, 47-61.

Yu HH, Rosa MGP (2010) A simple method for creating wide-field visual stimulus for electrophysiology: mapping and analyzing receptive fields using a hemispheric display. *J Vision*, *10*(14), 15-15.

Yu HH, Rosa MGP (2014). Uniformity and diversity of response properties of neurons in the primary visual cortex: selectivity for orientation, direction of motion, and stimulus size from center to far periphery. *Visual Neuroscience*, *31*(01), 85-98.

Yu HH, Verma R, Yang Y, Tibballs HA, Lui LL, Reser DH, Rosa MGP (2010) Spatial and temporal frequency tuning in striate cortex: functional uniformity and specializations related to receptive field eccentricity. *Euo J Neurosci*, *31*(6), 1043-1062.

Zavitz E, Yu HH, Rowe EG, Rosa MGP, Price NS (2016) Rapid adaptation induces persistent biases in population codes for visual motion. *J Neurosci*, *36*(16), 4579-4590.

Zeki SM (1974) Functional organisation of a visual area in the posterior bank of the superior temporal sulcus in the rhesus monkey. *J Physiol, 236*(3), 549-573.

Zeki SM (1991) Cerebral akinetopsia (Visual motion blindness). *Brain, 114*, 811-824.