Neural representation of complex motion in the primate cortex

Dissertation

for the award of the degree

"Doctor rerum naturalium"

of the Georg-August-Universität Göttingen

within the doctoral program Systems Neuroscience
appendant to the Göttingen Graduate Center for Neurosciences, Biophysics, and Molecular Biosciences (GGNB)

of the Georg-August University School of Science (GAUSS)

submitted by

Benedict Wild
from Regensburg

Göttingen, 2021

Thesis Committee

Prof. Dr. Stefan Treue (supervisor)
Cognitive Neuroscience Laboratory, German Primate Center

Prof. Dr. Tim Gollisch
Department of Ophthalmology, University Medical Center Göttingen

Prof. Dr. Michael Wibral
Department of Data-driven Analysis of Biological Networks, University of Göttingen

Members of the Examination Board

Referee: **Prof. Dr. Stefan Treue** (supervisor)
Cognitive Neuroscience Laboratory, German Primate Center

2nd referee: **Prof. Dr. Tim Gollisch**
Department of Ophthalmology, University Medical Center Göttingen

Further members of the Examination Board

Prof. Dr. Michael Wibral
Department of Data-driven Analysis of Biological Networks, University of Göttingen

Prof. Dr. Hansjörg Scherberger
Neurobiology Laboratory, German Primate Center

Prof. Dr. Tobias Moser
Institute for Auditory Neuroscience & InnerEarLab, University Medical Center Göttingen

Prof. Dr. Siegrid Löwel
Department of Systems Neuroscience, Institute for Zoology and Anthropology, University of Göttingen

Date of oral examination: May 28, 2021.

Contents

Acknowledgements

It is a truism that the work described in a PhD dissertation is never the result of just that one person's labor whose name is on the title page. Specific contributions to individual projects are mentioned in each chapter, but I want to use this space to thank people whose support went way beyond helpful comments or assistance during data collection.

I would like to thank Stefan Treue for giving me the opportunity to conduct my PhD work in his lab, for his guidance throughout the years, and for providing many opportunities for me to learn, such as attending summer schools and conferences.

A heartfelt thanks goes to Tim Gollisch and Michael Wibral for serving on my thesis committee and providing helpful advice. I also thank the remaining members of my examination board, Hansjörg Scherberger, Tobias Moser, and Siegrid Löwel, who have agreed to invest their time in reading my thesis, listening to my presentation, and providing thoughtful criticism.

Other people have provided additional mentorship to me without having their names mentioned in any official documents. I am very grateful to Alexander Gail and Igor Kagan for their support and feedback during my time at the DPZ.

I would like to thank my mentors and supervisors during my Bachelor's and Master's studies without whom I wouldn't even have made it to Göttingen in the first place: André Rupp at the University of Heidelberg, Paul Sauseng at the University of Munich, and Donatas Jonikaitis in Munich (now at Stanford University).

Two people fall into both categories, "Mentors" and "Friends and colleagues": I would like to thank Amr Maamoun for teaching me electrophysiology, for letting me join his project, and for guiding me throughout the first

few years of my PhD. Equally, my thanks goes out to Cliodhna Quigley with whom I have had the honor to run the psychophysics course for several years, during which I learned a lot from her about psychophysics, about teaching, and about science more broadly.

I have been very lucky in having had the chance to supervise four amazing Bachelor and Master students: Ann-Kristin Kenkel, Leonard Posner, Yifan Mayr, and Ilona Vieten. Thanks so much for putting up with me, challenging me, and for all your help with different projects.

Neurophysiology in nonhuman primates is definitely not a one-player game. I am very grateful to Isaac Guillen and Akshay Edathodathil for all their help and countless hours that we spent together in the setups.

Despite all the wonderful and hard-working scientists, the Cognitive Neuroscience Lab wouldn't last more than two days[1] without its core: a very, very special thanks goes to...
... Leo Burchhardt for teaching me handling and training of the animals, the patience associated with that, and for taking care of my monkeys whenever I went on vacation;
... Sina Plümer for helping me with everything: setup problems, monkey problems, protocol problems, life problems;
... Janine Kuntze for spending so much time training "monkey T";
... Luisa Klotz for making sure that I never ran out of chamber caps;
... Ira Panolias, Daniela Lazzarini, and Ronja Mielsch for all their help;
... Klaus Heisig for building anything I could ever ask for;
... Beatrix Glaser for all the administrative help and for being so patient with my extra requests and my inability to fill out business trip forms;
... Matthis Drolet for technical and administrative help, especially with teaching;
... Ralf Brockhausen for too many things as that I could mention all of them here (see chapters 4, 5, and 6), but I hope I can make his day by saying: Hurz!
... all of you for all the work you do and for countless (pre-Covid) breakfast sessions in the seminar room.

Similarly I am very much indebted to our animal caretakers Tanja Magerhans, Janine Kramer, Carina Hunger, Lea Borchard, Janine Senger, Suse

[1] Experiments have been conducted and it has been confirmed that somewhere around 30 h is the maximum that the lab can survive without TAs before chaos erupts.

Kurre-Passek, and especially Max Michel, who has been around from may first day until today and who is the only one at the DPZ who is interested in talking US sports. I am also very grateful to the DPZ's veterinarians, especially Tamara Becker, Annette Schrod, Karen Lampe, and Birgit Kamp, for all their efforts in ensuring our monkeys' well-being.

A shout-out to many wonderful colleagues and friends who have made life in the Cognitive Neuroscience Lab so much more fun: Lauren, Laura, Julia, Pinar, Antonino, Dana, Ralf, and many more. My co-founders of "NWG", "NCM", and "SfN" (you know what I'm talking about), Michael and Philipp, are wonderful co-workers, but – more importantly – great friends (and conference-roommates!), even though neither of them knows what makes a good soccer team. Thanks so much guys!

A special thanks goes to Michael Berger, Lauren Cassidy, Felix Schneider, Akshay Edathodathil, Laura Hansmeyer, Philipp Ulbrich, Julia Nowak, Bill Galic, and Pankhuri Saxena for proofreading and providing feedback on parts of this dissertation. I am also immensely grateful to Zurna Ahmed for all her help in designing the implants for my monkeys.

Doing a PhD in the natural sciences is a trying and demanding endeavor and it has been my experience that – in addition to great friends – a hobby is helpful in order to stay sane. I have found sanity (and sometimes more insanity) in running, which I had taken up almost exactly at the same time as when I joined the Cognitive Neuroscience Lab. Starting out with the idea of running a marathon just for the sake of crossing it off my bucket list, I have come to embrace what one of my coaches said: "Running is not a sport, it's a lifestyle". I would like to thank my teammates at LG Göttingen, as well as all the folks around Jay Stephenson and Ryan Hall at Run Free Training, especially Colin Schultz, who has guided me on my athletic journey for the last one and a half years.

I have had people who supported me as a scientist and I have had people who supported me as a runner. But the biggest thanks goes out to those who have supported me all the way in every aspect of my life: my "third parent" Bill Galic, for opening his home to me 15 years ago and changing my life forever; my Mom & Dad for all the opportunities they have provided me with, for never sparing any costs when it came to my education, and, most of all, for their endless love and support; and Pankhuri, for Lifting me up when I was down, for Offering support every day, for Valuing me as I am, and for Encouraging me in everything I do.

Declaration

Hereby I declare that I have written this thesis independently and with no other aids and sources than quoted.

Göttingen, March 31, 2021.

(Benedict Wild)

Chapter 1

Introduction

The brain has been described as "the most sophisticated and complex organ that nature has devised" (Bear et al., 2007) and even as the "most complex piece of active matter in the known universe" (Koch, 2020)[1]. A number of seemingly simple, biophysical processes, including the flow of ions through cell membranes (Hodgkin & Huxley, 1952) and chemical messengers that transmit signals between cells (Elliott, 1904; Loewi, 1921; Katz, 1969), give rise to all of our sensations, perceptions, thoughts, emotions, decisions, and actions. Running billions of these biophysical processes in parallel makes the brains of highly evolved animals capable of remarkable achievements. Especially when thinking of the human brain, a wide range of impressive mental performances come to mind: chess grandmasters who can remember hundreds of games and calculate dozens of possible outcomes of a series of moves; artists who write, compose, or paint masterpieces that move people to tears; scientists who attempt to "learn what, deep within it, binds the universe together"[2], often paving the way for revolutionary new technologies; or political leaders, who inspire their followers with grand ideas and great charisma.

And yet, all modern brains – human and nonhuman – were not designed to play chess, compose symphonies, or ponder the relation of mass and energy, but are simply the product of millions of years of natural selection that favors evolutionary fitness. And while the memory of a chess player or the creativity of an artist might improve one's chance of procreating in at least part of

[1]Statements such as the ones cited here typically refer to the human brain. While humans are capable of cognitive feats that exceed anything seen in the animal kingdom, the complexity of the monkey, cat, and even mouse brain is equally staggering. With that in mind, "brain" shall here refer mostly to the mammalian brain, in full acknowledgment that many other animals have similarly complex brains, the detailed treatment of which, however, exceeds the scope of this thesis.

[2]Goethe, Faust I, Vs. 382-383, translated by Stuart Atkins.

today's world, these highly specific skill sets are certainly not what guaranteed survival generations ago[3]. For most of history, reproductive success depended on the sheer ability to survive long enough to find a suitable mate[4]. Survival, in turn, depends on one's ability to approach and make use of beneficial stimuli (such as food), while avoiding harmful and dangerous ones (such as predators). The purpose of a nervous system, with the brain as its centerpiece, can thus be summarized as follows:

- Gather information about the environment by registering physical signals and transducing them into a form that can be processed by the nervous system (typically referred to as "sensation").
- Evaluate these signals as (potentially) beneficial or (potentially) harmful and decide on the appropriate course of action.
- Perform the selected action.

The first and the last of these three points have become two more or less self-contained subfields of research within neuroscience, commonly referred to as "sensory neuroscience" and "motor neuroscience"[5]. The second point is often referred to as "decision making", "cognitive control", "executive functions", or simply "cognition" and has a large overlap with at least one of the other two points[6]. Furthermore, the process of evaluating the sensory input must assign emotional valence (the degree to which something is pleasurable or aversive) to stimuli, which affects the organism's motivation to perform different actions (Tye, 2018). It should be obvious that a complete understanding of the brain will need to encompass and integrate all three of these points (e.g., Panzeri et al., 2017). And indeed, the last 10 years

[3]Of course, memory and creativity are mental capacities that have always been advantageous, and it is likely that such skills could have been put to use for other purposes in a society that does not know about chess or does not value art. I merely mean to point out that the specific ability to excel in chess or composing are not what brains have been optimized for by evolution.

[4]For most animals, this still holds true today.

[5]For example, two of the most widely used neuroscience textbooks dedicate large parts to "Perception" and "Movement" (Parts V and VI in Kandel et al., 2013) or "Sensory and Motor Systems" (Part II in Bear et al., 2007). Of the 10 "Themes" the *Society for Neuroscience* uses to structure its annual meeting, one is called "Sensory Systems" and one is called "Motor Systems".

[6]For example, "decision making" as it has been studied by the labs of Bill Newsome and Michael Shadlen (Britten et al., 1992, 1996; Shadlen & Newsome, 1996; Gold & Shadlen, 2007), is intimately related to motion perception, i.e., sensation. "Decision making", as it has been studied by the labs of Paul Cisek and John Kalaska, is intimately related to the planning of motor actions (Cisek & Kalaska, 2005, 2010; M. Wang et al., 2019).

have seen a shift towards a more holistic approach of studying the brain, which has been made possible by new technologies, such as two-photon Ca^{2+} imaging (Helmchen & Denk, 2005; Grewe et al., 2010; Allen et al., 2017) and high-density recordings with silicon probes (Jun et al., 2017; Steinmetz et al., 2019). These new approaches have clearly demonstrated that dividing the brain into independent regions that perform isolated tasks is an oversimplification. The activity of brain areas in rodents that were once considered to be purely sensory is, in fact, modulated by actions (Niell & Stryker, 2010; Saleem et al., 2013). High-density recordings have confirmed that neurons whose activity varies according to visual input, decision making, or action are distributed widely across the rodent brain (Steinmetz et al., 2019). However, these large-scale recordings also pose a problem: how are we to interpret brain-wide activity patterns without at least a rough reference frame to differentiate incoming (sensory) from outgoing (motor) signals? In other words, if every aspect of behavior and cognition is encoded by every part of the brain, how can we make sense of it all? The answer lies, as always, in the middle. Scientific progress is never achieved by one individual solution but relies on the collaboration of a large community with diverse approaches. Holistic, integrative methods and specialized, targeted experiments are not two opposing ideas, but two parallel paths that complement one another.

With this in mind, the work presented in this dissertation follows the second, more reductionistic way; not because I believe this to be the one true way forward, but because this focus on one small aspect of the brain can serve as a puzzle piece that contributes to the proverbial bigger picture. I have chosen "my puzzle piece" from the field of sensory neuroscience, more specifically, the perception of visual motion. As I highlight in the remainder of this introduction (and throughout the dissertation), this topic is highly important and serves well as a model for sensory processing and, indeed, neural processing more generally. The bulk of the dissertation (chapters 2-6) is concerned with how neurons in the medial superior temporal area (MST) – a high-level, extrastriate area of the primate visual cortex – represent information about visual motion. Additionally, a psychophysics experiment (chapter 7) investigates how different physical stimulus features affect motion perception in a way that suggests differential involvement of specific parts of the brain's motion processing pathway. To motivate this research, I will first briefly review some general topics of sensory neuroscience, before turning to vision as one of the best-studied sensory systems. After a brief interlude to discuss how information can actually be represented by neural activity, I will then describe how motion is processed in areas of the primate brain leading up to MST and why MST itself is an excellent model to study sensation, cognition, and neural processing.

1.1 Sensory neuroscience

As pointed out in the previous section, one of the fundamental purposes of the nervous system is to gather information about the environment. The goal is not to create the most accurate description of the physical world as it is, but to generate an internal representation of those aspects that are relevant for survival. A complete, perfectly accurate description of the world must necessarily contain all the relevant aspects (along with additional, less relevant bits and pieces) and might therefore seem desirable. However, one need to keep in mind that the biophysical processes that are required to generate this internal representation cost energy (Attwell & Laughlin, 2001; Lennie, 2003). Therefore, a more cost-effective, "sparse" representation of only the relevant elements of the environment is preferable (Olshausen & Field, 2004).

Of the external[7] senses that provide information about the world, the ones that offer the most direct evidence of harm or benefit are probably the somatosensory and the gustatory system: the sense of touch can immediately alert an organism to damage that is being done to the body (e.g., through extreme pressure or temperature) and the sense of taste offers valuable information about the nutritious value of food (e.g., sweetness is an indication for the presence of carbohydrates) or potential danger from poisons (which often have a bitter taste). However, both these systems suffer from a weakness: they require direct physical contact between the organism and the object of interest, which is highly dangerous if the object is harmful and only of limited use when the object is desirable.

In contrast, the visual, auditory, and olfactory systems provide information about objects in the external world without direct physical contact. Different species across the animal kingdom have optimized different sensory systems (or combinations of systems), depending on their ecological niche, to maximize their evolutionary fitness. Primates, both human and nonhuman, have adapted to deal with a variety of environments with very different requirements. Their dexterity allows them to physically interact with all kinds of objects and use them to their advantage, which gives a special role to vision for sensory input: whether it is plucking a fruit from a tree, climbing said tree while fleeing from a predator, or throwing a speer at a prey animal, none of these actions would be possible without accurate visual perception. And indeed, about 50% of the cerebral cortex in macaque monkeys and 20-30% of the human cortex are dedicated to the processing of visual information (Orban et al., 2004). The intuitive importance of vision – the loss of which probably

[7] I will forgo a discussion of interoception, which would exceed the scope of this dissertation.

affects our everyday lives more than the loss of any other sense – has led generations of neuroscientists to explore this sensory system in more detail than any other.

1.2 Vision

The ability to see seems like the most natural thing in the world: "We just open our eyes and look!" (Palmer, 1999). However, problems quickly arise when thinking about the details of visual perception. The first question is how what we see in the external world becomes accessible to the brain.

The physical signal that serves as the basis of vision is light that is reflected by objects in the world. This reflected light enters the eye through the cornea, pupil, lens, and vitreous humor before it has its first contact with the nervous system, the *retina*. The photoreceptor cells of the retina contain transmembrane proteins, so-called *opsins*, which react to photons by means of the chromophore *retinal* (Wald, 1968). Two types of photoreceptors, *rods* and *cones*, transmit graded potentials on to *bipolar cells* which, in turn, pass the signal on to *retinal ganglion cells* (RGCs). Two additional cell types, *horizontal cells* and *amacrine cells*, modulate the signal transmission from photoreceptors to bipolar cells and from bipolar cells to RGCs, respectively. A complex network of excitatory and inhibitory synapses between these six classes of cells – most of which have multiple types – distributes the light signal into multiple parallel pathways (Wässle, 2004) and performs an impressive amount of signal processing, such as temporal filtering and thresholding (Gollisch & Meister, 2010). The axons of RGCs make up the *optic nerve* which, in mammals with forward-facing eyes, transmits the output of the retina to the *optic chiasm*. Here the optic nerve fibers coming from the medial (nasal) side of each retina cross over to the other hemisphere while those fibers originating from the lateral (temporal) part of the retina stay within their respective hemisphere. These merged bundles of axons relaying information from one half of the visual field to the contralateral hemisphere continue as the *optic tract* to the *lateral geniculate nucleus* (LGN) of the thalamus. Neurons in the primate LGN are arranged in 6 layers that each receive information about the contralateral hemifield from one eye. These signals are then passed on through the *optic radiation* to the equally six-layered *striate* or *primary visual cortex* (V1) (Henderickson et al., 1978). From here on the signals that have so far been transported by tightly packed, parallel fibers through a single pathway[8] are split into spatially separated

[8]In addition to the *geniculostriate pathway* described here, a second, *retinotectal pathway* sends part of the optic tract to the superior colliculus (May, 2006; Wilson & Toyne, 1970),

pathways consisting of dozens of specialized brain regions (Felleman & Van Essen, 1991). The most important organizational scheme in the primate visual cortex is that of two distinct processing streams: a ventral pathway, including area V4 and the inferior temporal (IT) cortex, that is predominantly concerned with visual recognition of objects, including features such as orientation and color; and a dorsal pathway, including the middle temporal (MT) and medial superior temporal (MST) areas, that is involved in the perception of spatial relationships, including motion (Mishkin et al., 1983; Maunsell & Newsome, 1987; Goodale & Milner, 1992) (Figure 1.1).

The coarse anatomical structure of the visual system that has been described until now suggests a processing pipeline in which a series of stages – RGCs, LGN, V1, higher cortical regions – transform the visual signal that was received by the retinae to extract behaviorally relevant information. The following paragraphs highlight some of these processing steps, as they provide the basis for the results described in subsequent chapters.

RGCs are among the best studied cells of the brain. Hartline (1938) first noted that RGCs can be classified based on how they respond to changes in the conditions of illumination. "On" cells respond with an initial burst of action potentials when light is turned on and then continue to fire at a steady, but much lower rate; "off" cells show the opposite pattern, where no action potentials are fired during illumination, but they discharge vigorously when the light is turned off; "on-off" cells, at last, respond to any change (light on or light off) with a brief burst of action potentials, but remain mostly quiet during steady levels of illumination. Hartline (1938) also introduced the term *receptive field* to visual neuroscience[9] to describe "the region of the retina which must be illuminated in order to obtain a response". Today this is often used to refer to the part of the visual field (rather than the retina) that can modulate a visual neuron's activity. The receptive fields of RGCs were explored in more detail by Kuffler (1953), who found that they typically had an antagonistic structure: stimulation of one part of the receptive field led to a strong "on" response (as described above), whereas stimulation of other parts of the receptive field led to a strong "off" response. In many RGCs, these "on" and "off" regions were arranged in a concentric manner with an excitatory (or inhibitory) area in the center of the receptive field and an inhibitory (or excitatory) area surrounding it (see also Wiesel, 1959).

A different way of classifying RGCs is based on their morphology: small *midget cells* with compact dendritic arbors, and larger *parasol cells* form

a midbrain structure that is involved in eye movements (Gandhi & Katnani, 2011) and will not be discussed here.

[9]The term had previously been used by Sherrington (1906) to describe an area of skin that, when stimulated, will lead a neuron to respond.

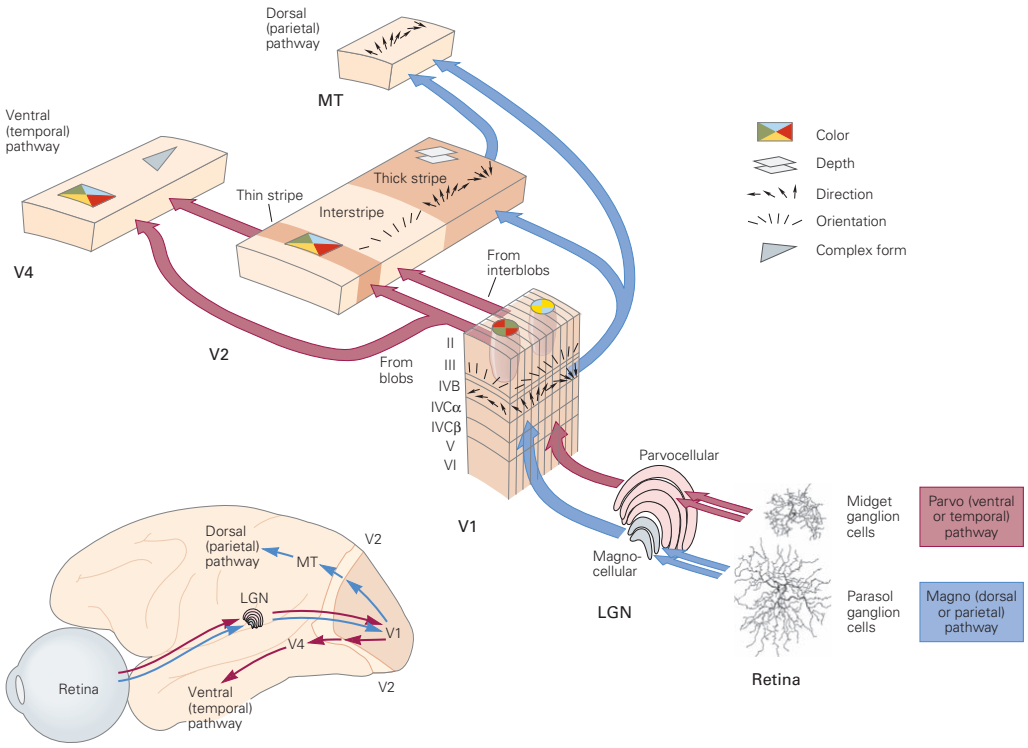

Figure 1.1: The visual processing pathway: light enters the eye and is transduced into electrical signals by photoreceptors in the *retina*. Two types of retinal ganglion cells, *midget cells* and *parasol cells*, pass the signal on to the *parvocellular* and *magnocellular* layers of the *lateral geniculate nucleus* (LGN). These two pathways and in layers 4Cβ and I4Cα of the *primary visual cortex* (V1). After V1 and V2, the processing pathway is split into the *ventral pathway* that includes V4 and the *dorsal pathway*, which includes MT (reprinted, with permission, from Kandel et al., 2013).

the two main categories, but many more anatomically distinct cell types exist (Callaway, 2005). Two classes of RGCs that vary in their contrast sensitivity have been shown to project to the four *parvocellular* (P) layers and to the two *magnocellular* (M) layers of the LGN (Kaplan & Shapley, 1986), and the evidence suggests that these two classes are in fact identical (or at least largely overlapping) with the midget and parasol cells, respectively (Callaway, 2005, Fig. 1.1)[10]. Neurons in the P and M layers of the LGN differ not only in their contrast sensitivity (low for P cells, high for M cells), but also in their receptive field size (smaller for P cells, larger for M cells) and in their sensitivity for color (color-opponent receptive fields in P cells, achromatic receptive fields in M cells) (Wiesel & Hubel, 1966). A third class of LGN neurons, *koniocellular* or K neurons (not shown in Fig. 1.1), lie within the intralaminar zone between P and M layers, but have received less attention than the M and P pathway (Casagrande, 1994; Callaway, 1998). All of this shows a surprising degree of segregation in how visual features are represented, even at the earliest stages of visual processing.

As the primary visual cortex (V1) receives input from the LGN, it needs to keep track of three pairs of pathways by which signals can arrive: from the left or the right LGN (which represents information about the right and left visual hemifields respectively); from the parvo- or magnocellular layers of the LGN; and from the left or right eye.

The first distinction is straightforward, as the left and right LGN project to V1 of the left and right hemisphere, thus passing their representation of the contralateral hemifield on to the visual cortex. Importantly, this representation of the contralateral hemifield is not random, but *retinotopically* organized, meaning that neighboring cells in the brain typically have neighboring, or even overlapping receptive fields.

Projections from the parvo- and magnocellular layers end in separate layers of V1, namely layers 4Cβ and 4Cα (Fig. 1.1). Layer 4C sends its main projections to layers 2-4B which are connected by a feedback circuit with layer 5, but also project to extrastriate areas, such as V2 and MT (Fig.1.1; see Callaway, 1998, for a review). The idea that two distinct pathways that are specialized for color/form (parvocellular layer of LGN – layer 4Cβ – superficial V1 layers – ventral cortical pathway) and motion (magnocellular layer of LGN – layer 4Cα – layer 4B of V1 – dorsal cortical pathway) was further supported by the finding that inactivation of the magnocellular, but not the parvocellular, layers of the LGN strongly affects responses in extrastriate area MT (part of

[10]One cannot ignore the irony in midget cells being associated with the parvocellular pathway and parasol cells being associated with the magnocellular pathway, a choice of terminology that must have driven generations of neuroscience students mad.

the dorsal pathway) (Maunsell et al., 1990). However, later research showed that the superficial layers of V1, which constitute the beginning of the ventral cortical pathway, receive substantial contributions from both parvo- and magnocellular pathways, suggesting that a strict segregation all the way from the retina to extrastriate cortex is oversimplified (Nealey & Maunsell, 1994).

The third distinction – which eye the signal is coming from – was addressed as part of the seminal work by Hubel and Wiesel. They found that neurons in the primary visual cortex of both cats (Hubel & Wiesel, 1962) and primates (Hubel & Wiesel, 1968) varied in how much their responses were dominated by stimulation of one eye or the other, a feature referred to as *ocular dominance*. Moreover, they found that cells with similar ocular dominance were clustered together in vertical slices through the six layers of the striate cortex that are now known as *ocular dominance columns*.

In addition to layers and ocular dominance columns, Hubel and Wiesel also discovered what is today the most widely known and investigated feature of the primary visual cortex: V1 neurons respond selectively to the orientation of a bar of light or a luminance edge. Whereas the receptive fields of RGCs and LGN neurons are well described by a concentric circular arrangement of an excitatory ("on") or inhibitory ("off") center and an antagonistic surround (see above), the receptive fields of many V1 neurons have excitatory or inhibitory regions that are elongated and flanked by antagonistic regions and can be oriented vertically, horizontally, or obliquely (Hubel & Wiesel, 1959). While this might seem mundane at first glance, the sharp emergence of selectivity for a feature that was absent in the input from the LGN has become the quintessential example of cortical computation (see Priebe & Ferster, 2012, and Priebe, 2016, for reviews). Hubel & Wiesel (1962) themselves suggested a basic model, in which a number of LGN cells with appropriately aligned receptive fields project to a V1 neuron that is then selective for the orientation of this alignment (Fig. 1.2). Such a model was supported by simultaneous thalamic and cortical recordings in cats that showed monosynaptic connections between LGN and V1 neurons with overlapping receptive fields (Tanaka, 1983; Reid & Alonso, 1995).

What has been fiercely debated in the decades after this model was proposed is whether the input coming from the LGN is sufficient by itself to generate orientation selectivity, or whether additional mechanisms within the circuits of V1 are necessary. On the one hand, applying an antagonist to the inhibitory neurotransmitter GABA broadens V1 neurons' tuning curves (Sillito, 1975), suggesting that intracortical cross-orientation inhibition plays an important role in sharpening orientation tuning. On the other hand, recordings of synaptic potentials in V1 neurons while inactivating the cortex by cooling it, showed that orientation tuning was mostly unaffected (Ferster

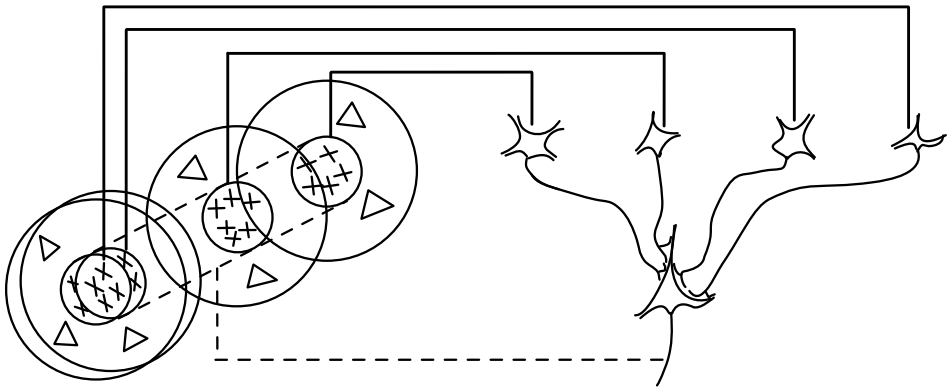

Figure 1.2: Proposed scheme for explaining the emergence of orientation selectivity in V1, based on input from the LGN. A large number of LGN cells, of which four are illustrated in the upper right in the figure, have receptive fields with "on" centers arranged along a straight line on the retina. All of these project upon a single V1 cell with excitatory synpases. The receptive field of the V1 cell will then have an elongated "on" center indicated by the interrupted lines in the receptive-field diagram to the left of the figure. Crosses indicate excitatory regions, while triangles indicate inhibitory regions of the receptive fields (reprinted, with permission, from Hubel & Wiesel, 1962).

et al., 1996), suggesting that the LGN input, that was left intact by the cooling, is sufficient to generate this selectivity. A third possibility is that LGN cells themselves show an orientation bias which is then sharpened by non-specific inhibition (Vidyasagar et al., 1996; Viswanathan et al., 2011). These conflicting explanations show that even for a seemingly simple and "well-behaved" part of the visual pathway, such as V1, the question of how receptive field structure and tuning arise is not yet fully understood. As later chapters will demonstrate, characterizing neural response functions only becomes more difficult in higher visual areas.

Whatever the exact mechanisms underlying orientation selectivity may be, the emergence of more complex selectivities from unselective input has become an essential building block for models of cortical computation (Riesenhuber & Poggio, 1999) and will come up again in later chapters of this dissertation (compare, for example, Fig. 1.2 with Fig. 2.5). As a matter of fact, the very next stage in the cortical processing pipeline that adds another layer of abstraction was also discovered by Hubel & Wiesel (1962)[11]: while most

[11]Given the modern obsession with quantifying research output by means of impact

cells in V1 are orientation selective, not all of them respond in the same way to oriented bars and edges. So-called *simple cells* have receptive fields with clearly separated "on" and "off" regions and their responses to visual stimuli can be predicted from a convolution of this "on/off" arrangement with the stimulus. *Complex cells*, on the other hand, show position invariant responses to luminance edges, meaning that they respond equally well to their preferred direction anywhere in the receptive field, irrespective of the exact arrangement of light and dark areas. This means that, unlike simple cells, they cannot be described by a single linear filter, instead relying on multiple filters that need to be combined (Rust et al., 2005; Vintch et al., 2015). Whether simple and complex cells really form two distinct populations (Skottun et al., 1991) or rather the two ends of a continuum (Mechler & Ringach, 2002) has been debated; but either way the representation within the primary visual cortex quickly shifts to a more abstract level. Both simple and complex cells respond more strongly to moving than to stationary stimuli, and about half of the complex cells in primate V1 appear to be *direction selective* in that they respond better to movement in one direction than in the opposite direction (Hubel & Wiesel, 1968).

Before moving to the topic of visual motion in more detail, I would like to take a brief detour to address an issue that has so far been touched upon in a rather informal manner: the question of how information about anything in the external world is actually encoded in neural activity.

1.3 The neural code

The previous sections made frequent use of the words "signal" and "information" that is being transmitted by a signal, without discussing these terms further. While it is quite difficult to find meaningful, non-circular definitions of these words, I loosely define *signal* as any change in a physical quantity that conveys information, and *information* as anything that reduces uncertainty[12]. The physical quantity that constantly changes in the brain is the distribution of ions inside and outside of individual neurons, which can be measured as an electric potential difference across cell membranes. These voltage differences come in two forms: as a continuous, oscillating voltage signal that can be measured intracellularly ("*membrane voltage*") or extracellularly ("*local field*

factors, citations, H-indices, and publication counts, it seems mind-boggling that Hubel & Wiesel published their discoveries of ocular dominance, the simple/complex cell distinction, and their model of orientation selectivity in a single paper.

[12]The difficulty in defining these terms quickly becomes obvious when trying to define "uncertainty" without using the words signal and information.

potential", LFP); and as a series of all-or-nothing events known as *action potentials* or *spikes*, that occur when the voltage signal crosses a threshold (see Hodgkin & Huxley, 1952, for biophysical details). By isolating large parts of neuronal axons with myelin, the conduction velocity of action potentials can be increased drastically ("saltatory conduction"), which makes them a highly efficient signal for rapid information transmission. Spikes, which have been described as the "neural code" (Rieke et al., 1999), can be recorded from individual or small clusters of neurons in living and even awake mammals. How information is encoded in *spike trains* (i.e., sequences of spikes in a finite time window) is one of the central questions in neuroscience. The fact that one and the same spike train can be associated very different experiences has been described as early as the 19th century:

> *Denn dieselbe Ursache kann auf alle Sinnesorgane zugleich einwirken, wie die Elektricität; alle sind dafür empfänglich, und dennoch empfindet jeder Sinnesnerve diese Ursache auf eine andere Art; der eine Nerve sieht davon Licht, der andere hört davon einen Ton, der andere riecht, der andere schmeckt die Elektricität, der andere empfindet sie als Schmerz und Schlag*[13] (Müller, 1838).

Today, this is referred to as *labeled line coding*: the reason why one train of spikes leads to the percept of a tone, while another, maybe even similar train of spikes leads to the percept of the color red, is that the first spike train comes from an auditory neuron that receives its input from hair cells in the cochlea, while the second spike train comes from a visual neuron that receives its input from photoreceptors in the retina. Similarly, within one sensory modality, the reason why one spike train signals the color red (or upward motion, or a square), while another, maybe similar spike train signals the color green (or downward motion, or a circle) is because the first spike train comes from a red-sensitive neuron that receives its input from cone photoreceptors that are selective for a specific wave length of light, while the second spike train comes from a green-sensitive neuron that receives its input from cone photoreceptors that are selective for a different wave length. Useful information about the environment can then, of course, only be represented by a population of neurons that are specialized for certain stimulus attributes[14].

[13] The same cause, such as electricity, can simultaneously affect all sensory organs, since they are all sensitive to it; and yet, every sensory nerve reacts to it differently; one nerve perceives it as light, another hears its sound, another one smells it; another tastes the electricity, and another one feels it as pain and shock. (Translated by Edwin Clarke and Charles Donald O'Malley)

[14] Note that neither this "labeled line" code, nor any of the other coding schemes described here actually explain how (or why) exactly a series of action potentials in

Labeled line coding, however, does not answer the question of what features of a spike train convey information. Given that a spike train is nothing more than a sequence of all-or-nothing events that occur within a finite time window, there are really only two measures that can vary: the number of spikes within the time window and their exact timing (or, equivalently, the periods of silence between spikes, the so-called *interspike intervals*, ISI).

The idea that information about features of a stimulus is carried by the number of spikes a neuron fires in a certain time window is known as *rate coding*. Like the labeled line code, rate coding has been known for a long time:

> *The message which a nerve fibre can transmit must consist of one or more discrete impulses and a continuous transmission of the excited state is impossible. In fact the only way in which the message can be made to vary at all is by a variation in the total number of the impulses and in the frequency with which they recur* (Adrian, 1928).

Rate coding is probably the best-explored neural code and forms the basis for a number of frameworks that attempt to characterize the relation between neural activity and external stimuli. The obvious problem of rate coding is that the counting of multiple spikes (e.g., by a downstream neuron) requires integration over time. Depending on the duration of the time window in which spikes are counted, this could slow down the processing of information. One way to address this problem is to use the response of several neurons with similar response properties in a short time window as an estimate of the firing rate (Shadlen, 2006).

Another solution to this problem is offered by the other major coding scheme, commonly referred to as *temporal coding*, in which information is conveyed through the exact timing of spikes. Because the measurement of exact spike times requires more advanced equipment, this idea has only been explored recently. In salamander retinal ganglion cells, for example, it has been shown that the latency of the first spike in response to a briefly presented stimulus varies across different stimuli (Gollisch & Meister, 2008). Recordings from the cat LGN have shown that spike time precision depends on the timescale of the stimulus, and information-theoretic approaches suggest that "relative", though not necessarily "absolute" precision (with regard to the stimulus) represents stimulus attributes (Butts et al., 2007). Even in an

a number of neurons lead to subjective experiences or *qualia*. This "hard problem of consciousness" (Chalmers, 1995) falls outside the scope of neuroscience and into the realm of philosophy (Dennett, 1991; Searle, 2004).

extrastriate area of the macaque visual cortex, such as MT, neurons have been shown to respond with very high temporal precision to repeated presentations of the same stimulus (Bair & Koch, 1996).

Whether neurons act as "integrate-and-fire" devices that rely on the number of spikes in a given time window to transmit information, or whether the exact timing of spikes and the ISIs define the neural code, has been the topic of a long and ongoing debate.

In a classic study, Softky & Koch (1993) analyzed the variability of spike trains recorded from primate visual cortex and compared the data to biophysically plausible models of how pyramidal cells integrate synaptic input. They found that, given the large number of synaptic inputs typically observed in the brain, an integrate-and-fire mechanism would even out all irregularities of the input and lead to a highly regular, nearly periodic output. In contrast, the actual data showed highly irregular output, which led to the conclusion that these neurons work as "coincidence detectors" that only fire when many postsynaptic potentials coincide at the millisecond scale and that precise spike times and the ISIs carry meaningful information (Softky & Koch, 1993). Shadlen & Newsome (1994) countered that this approach ignored the role of inhibitory postsynaptic potentials, which, when balanced with excitatory potentials, can lead to the highly irregular spike trains typically observed. Furthermore they argued that neuronal membrane time constants on the order of 8 to 20 ms speak strongly against coincidence detection at the millisecond timescale.

The distinction between rate coding and temporal coding becomes blurred, and even downright paradoxical, when describing a neuron's firing rate not as the actual number of spikes in a predefined time window (the "spike-count rate"), but as is often done, as a time-dependent probability function of a spike occurring in a very small time window (typically written as $r(t)$, Dayan & Abbott, 2001; Brette, 2015). The paradox was beautifully shown by Rieke and colleagues (1999) by using the phase-locking of a neuron's time-dependent firing rate to a periodic stimulus as an example (such as auditory neurons in response to low-frequency, pure tone, e.g. Rose et al., 1967). In such cases, the firing rate follows the oscillation of the stimulus (typically with a short lag), which means that spike times tend to occur at a specific phase of the stimulus. Therefore, the timing of spikes provides important evidence about a feature of the stimulus (namely its phase), even though this example was supposed to show that the firing *rate* is modulated by the stimulus. As Rieke et al. (1999) point out, this paradox stems from the fact that "firing rate" is not as easily defined as one would think, and always depends on a time window in which spikes are counted or for which the probability of a spike is expressed. Making this time window arbitrarily small brings a rate code

closer and closer to a temporal code (Rieke et al., 1999; DeCharms & Zador, 2000).

Thus, the apparent dichotomy between rate coding and temporal coding is not as clear as it might seem at first. Evidence that information about a stimulus can be decoded from either exact temporal patterns of spikes or from the number of spikes in a fixed time window does not mean that the other coding scheme is useless. Some more recent examples in the primate somatosensory cortex even suggest that rate and temporal codes can both be used to represent different aspects of a stimulus (Harvey, Saal, Dammann, & Bensmaia, 2013; see also Zuo et al., 2015 for a similar example from rat somatosensory cortex).

More recently, it has been argued that "coding" (in the information theoretic sense of whether firing rates provide any information about a stimulus) is misleading, as it makes no statement about the causal role of the spike train in processing the stimulus (Brette, 2015, 2019). Specifically, the firing rate (in the probabilistic sense that is calculated based on data from multiple trials) of a presynaptic neuron cannot have a direct effect on a postsynaptic neuron because the postsynaptic neuron only has access to a single spike train, which is the realization of a random point process that is based on the presynaptic neuron's firing rate. While this is true, of course, I would argue that this does not invalidate the usefulness of firing rates for trying to understand neural processing. This is similar to arguing against any attempt of relating the activity of V1 neurons to a visual stimulus on the grounds that V1 neurons have no direct access to the actual stimulus, but only indirect access through spike trains coming in from the retina and the LGN. This is also true, but it does not mean that one cannot learn about cortical computation by observing how a V1 neuron's activity is modulated systematically due to changes in the stimulus. Similarly, a presynaptic neuron's firing rate clearly has an indirect effect on a postsynaptic neuron, in that the actual spike train that is instantiated at any given time depends (stochastically) on the firing rate. The trial-averaged firing rate is an abstraction and a simplification, but a useful one, and a large part of what is known about the brain today is the result of attempts to relate firing rates to sensory stimuli and behavior.

When it comes to deciphering the neural code, two broad classes of approaches have been used. What I call the "forward" approach presents the nervous system with a specific stimulus and asks "How does a neuron (or multiple neurons) respond to this stimulus?" This approach can be used to determine a neuron's response to repeated presentations of a stimulus. If the response varies systematically with a feature of the stimulus, the neuron is considered to be *tuned* for this feature. Famous examples for such a

tuning include the aforementioned orientation tuning in V1 cells (Hubel & Wiesel, 1962) and the tuning for the direction of arm movements in motor cortex (Georgopoulos et al., 1986).

Alternatively, what I call the "reverse" approach records the spiking responses to noisy, random input with particular statistical properties and asks "What features of the stimulus make the neuron fire an action potential?" These approaches make less assumptions about the relevant stimulus features that drive a neuron's activity. Instead they attempt to find these relevant features by exploring how stimuli that elicited a response (the so-called "spike-triggered stimulus ensemble", STE) differ from those that did not elicit any response. For example, a neuron could spike whenever one or several stimulus features deviate from the average of those features across all stimuli. In that case the STE would be shifted in the high-dimensional stimulus space with respect to the raw stimulus ensemble. The difference in the mean of the STE and the mean of the raw stimulus ensemble is called the *spike-triggered average* (STA). If the underlying stimulus distribution is radially symmetric and if the neuron's response function to this stimulus ensemble can be described by a linear filter (the *spatiotemporal receptive field*), possibly with the addition of a nonlinearity, the STA provides an unbiased estimate of this filter (Chichilnisky, 2001). Additional filters can be detected by investigating whether the STE differs from the raw stimulus ensemble in terms of the relation between different stimulus features. Such changes in the variance of the STE can be explored through *spike-triggered covariance* (STC) analysis (Rust et al., 2005; Touryan et al., 2002). Together, these methods have allowed to describe the functional relationship between external stimuli and neuronal responses in much more detail and complexity than one-dimensional tuning curves, albeit at the cost of much more data that is required to fit such models (see Schwartz et al., 2006 and Sharpee, 2013, for reviews).

Having established the elementary principles of how visual information is processed in the primate brain, how information can be encoded in spiking activity, and how this neural code can be deciphered, I will now address that subset of visual neuroscience that is the focus of this dissertation: the representation and processing of visual motion.

1.4 Motion perception

Our world is highly dynamic, which is best summarized by the famous phrase attributed to Heraclitus:

"The only constant is change."

Motion, defined as a change in position over time, is ubiquitous and an essential part of our environment. A famous case study of motion blindness (or *akinetopsia*) is that of patient L.M., who suffered bilateral damage in the posterior brain. This case illustrates how severely one's everyday life is affected by the inability to see how objects in the environment change their position: "She had difficulty, for example, in pouring tea or coffee into a cup because the fluid appeared to be frozen, like a glacier. In addition, she could not stop pouring at the right time since she was unable to perceive the movement in the cup (or a pot) when the fluid rose" (Zihl et al., 1983). As illustrated by this example, perception of motion, more than shape or color, is intimately related to action: when objects of interest move, this movement often requires some sort of behavioral reaction, such as moving one's hand to stop pouring tea. Eye movements in particular are important to follow behaviorally relevant moving targets. The neural circuits that underlie motion perception and those that are involved in the planning and execution of eye movements therefore form a "microcosm" of the perception-cognition-action loop that was outlined previously. In addition to its relevance for everyday life and direct connection to action outcomes, motion can also be easily defined with two parameters that provide all of the essential information that needs to be encoded: direction and speed. All of these points make the processing of visual motion an ideal model for neural computation.

From the visual system's point of view, motion means that light that is reflected by one and the same stimulus falls on different photoreceptors at different points in time. Thus, a comparison has to be made in both space and time. The spatial comparison can be done by comparing the activity of cells with neighboring receptive fields. To account for the temporal aspect ("Which cell was activated first?"), a short delay needs to be incorporated into the comparison. A detailed implementation of such a comparison has first been proposed by Hassenstein & Reichardt (1956) (Fig. 1.3A).

Whether and how such "Reichardt detectors" are implemented in the nervous system has been investigated in both the mammalian retina, inspired by the description of direction-selective RGCs in the rabbit (Barlow & Hill, 1963), and in the optic lobe of flies (see Mauss, Vlasits, Borst, & Feller, 2017, for a review).

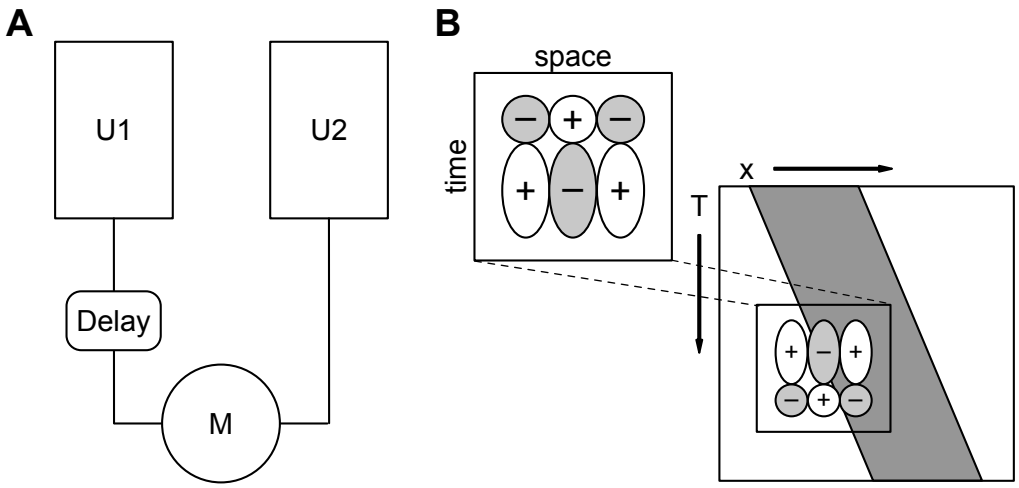

Figure 1.3: Models for motion detection. ***A***. Reichardt-Detector. The output of two units (U1 and U2) is multiplicatively combined by unit M, but the signal from unit U1 is delayed. Therefore, M will be maximally activated if U2 is activated shortly after U1. ***B***. Motion Energy Model, adapted from Adelson & Bergen (1985). A vertical bar moving from left to right appears as a slanted strip in an $x - t$ plot where one spatial dimension (horizontal axis) and one temporal dimension (vertical axis) are shown (right side of panel B). A fixed spatiotemporal filter (left side of panel B) that is convolved with a spatiotemporal stimulus that moves beneath it will respond strongly when motion lies within its receptive field. Panel B adapted with permission from Adelson & Bergen (1985), © The Optical Society

A seemingly alternative framework for the emergence of direction selectivity was proposed by Adelson & Bergen (Adelson & Bergen, 1985). Their "Motion Energy" model describes motion as an orientation in spacetime (with, for ease of visualization, a single spatial dimension), similar to normal orientation in ordinary 2D space. Analogous to spatial orientation selectivity described by Hubel & Wiesel (Hubel & Wiesel, 1962), a cell with spatiotemporal orientation selectivity, or, in other words, a cell with a spatiotemporally separable impulse response function, would be motion sensitive (Fig. 1.3B). Spatiotemporal impulse functions are frequently observed in visual neurons (e.g., DeAngelis et al., 1993a, 1993b), and such units can be described as filters of spatiotemporal energy. The model as it is presented in Fig. 1.3B is only motion sensitive, but not direction selective (i.e., it cannot differentiate between leftward and rightward motion). Direction selectivity can easily be added by replacing the impulse response with a Gabor function that is ori-

ented in spacetime. Because such filters are still flawed in that they are phase sensitive, the final part of the Motion Energy model consists of summing the squared output of two spatiotemporal filters that differ only in their phase. Although the Reichardt-detector and the motion energy model initially seem to differ, Adelson & Bergen (1985) pointed out that it is possible to construct motion energy models that are functionally equivalent to Reichardt detectors.

The similarity of the the spatiotemporally oriented filters in the Motion Energy model with the spatially oriented receptive fields described by Hubel & Wiesel (1962, 1968) is no coincidence. In primates, contrary to other mammals (Barlow & Hill, 1963; Taylor & Vaney, 2003), direction selective responses are first observed in V1 (Hubel & Wiesel, 1968). However, these cells suffer from the "aperture problem": they can only detect motion in a direction that is orthogonal to any edge or border that extends beyond its receptive field (Adelson & Movshon, 1982; Movshon et al., 1985; Simoncelli & Heeger, 1998; Pack & Born, 2001). A more general representation of velocity, that is less dependent on the exact physical features of the visual input, emerges in the extrastriate area MT (also known as V5). MT, located on the posterior bank of the superior temporal sulcus, was first described by Allman & Kaas (1971) in the owl monkey (*Aotus trivirgatus*), who coined the name "MT", and by Dubner & Zeki (1971) in the rhesus macaque (*Macaca mulatta*), who described the direction selective responses. It receives most of its input from V1 layer 4b (Figure 1.1) and additional inputs from V1 layer 6, V2, and V3 (Maunsell & Van Essen, 1983a). Neurons in MT are retinotopically organized (Van Essen et al., 1981) and also show columnar organization, similar to V1, based on preferred direction (Zeki, 1974; Albright et al., 1984) and preferred binocular disparity (DeAngelis & Newsome, 1999). MT neurons' tuning for motion direction has been confirmed by a large number of studies in different primate species (Zeki, 1974; Van Essen et al., 1981; Maunsell & Van Essen, 1983c; Albright, 1984; Felleman & Kaas, 1984; Snowden et al., 1992; Movshon & Newsome, 1996; Treue & Andersen, 1996; Pack & Born, 2001; Pack, Berezovskii, & Born, 2001; Rust et al., 2006; Wallisch & Movshon, 2019). Furthermore, MT neurons are also tuned for speed (Maunsell & Van Essen, 1983c; Lagae et al., 1993; Perrone & Thiele, 2001, 2002; Priebe et al., 2003), for binocular disparity (Maunsell & Van Essen, 1983b; DeAngelis & Newsome, 1999; DeAngelis & Uka, 2003), and for stimulus size by means of the suppressive surround structure of their receptive fields (Raiguel et al., 1995; Born, 2000).

A series of studies by the Newsome lab have provided evidence that the selective responses for motion direction in MT are strongly linked to motion perception. Consequently, this evidence has established MT as a prominent model for visual processing, perceptual decision making, and the relation of

neural activity and behavior. The task that is used in many of these studies presents a monkey with a random dot pattern (RDP) in which a certain percentage of dots move coherently in one direction, while the remaining dots move in random directions. The monkey signals which of two possible directions it perceived by making an eye movement to one of two choice targets that are associated with the two possible directions. The difficulty of the task and the monkey's performance depend on the strength of the motion signal, which is defined by the percentage of coherently moving dots (or simply *coherence*). Evidence that MT is necessary to perform this kind of task first came from a lesion study, which showed that motion thresholds (i.e., the coherence level that is necessary for the monkey to complete the task with a certain performance level) were elevated after lesioning MT with injections of ibotenic acid (Newsome & Paré, 1988). Microstimulation of a small patch of MT with current pulses of 10 μA biases a monkey's motion perception towards the preferred direction of the stimulated neurons (Salzman et al., 1990, 1992), suggesting that changes in MT activity are also sufficient to modulate perception. To further quantify the relation between neuronal activity and behavior, Newsome and colleagues developed the concept of a *neurometric function* (Newsome et al., 1989; Britten et al., 1992): They measured an MT neuron's response to RDPs of varying coherence levels moving in the neuron's preferred or anti-preferred direction. Then they calculated a receiver operating characteristic (ROC) curve by plotting for a range of threshold values (e.g., 0 - 120 spikes/s) the probability that a neuron's response to its preferred direction exceeds each threshold versus the probability that the response to the anti-preferred direction exceeds the same threshold. The area underneath this curve ("area under the ROC curve", auROC), which falls between 0.5 and 1, can be considered the probability of the neuron correctly identifying the presented motion. Plotting these auROC values for a range of coherences results in a "neurometric" function, which can be compared to the monkey's psychometric function (i.e., the probability of correctly identifying the direction of motion as a function of coherence). The authors found that direction discrimination thresholds of individual neurons were surprisingly similar to the behavioral threshold of the monkey and sometimes even better (Newsome et al., 1989). This might be explained by the fact that the analysis is based on single neuron responses to stimuli that were optimized to drive these very neurons, whereas in reality decisions have to be made based on the pooled output of many neurons with heterogeneous response properties (Britten et al., 1992). While this line of research showed that MT neurons represent motion direction with great sensitivity, it does not provide any information about how this representation is related to behavior. To address this question, another measure was developed, based on

the ROC approach outlined above, that specifies the probability that an ideal observer could tell, based on the firing rates from two trials, on which trial the monkey made which choice (Britten et al., 1996). This *choice probability* can theoretically vary between 0.5 and 1 (with 0.5 meaning chance performance and 1 meaning perfect prediction of behavior based on neural responses), with typical average values for sensory areas falling in the range of 0.52 (Gu et al., 2008) to 0.67 (Dodd et al., 2001) (see Crapse & Basso, 2015, for a review).

How does the motion selectivity of MT neurons arise? Based on the Motion Energy model, Simoncelli & Heeger (1998) proposed a two-stage model of motion processing in areas V1 and MT of the primate cortex that can explain a series of psychophysical and electrophysiological findings. The V1 stage contains both simple cells, which are rectified and normalized linear filters of the input image, and complex cells, which pool over simple cells with identical spatiotemporal orientation in a small spatial region, similar to the Adelson & Bergen (1985) model. In the second, MT-like stage, the output of direction-selective V1 complex cells is combined to construct an "intersection-of-constraints" scheme, which can solve the aperture problem that V1 cells are faced with (Adelson & Movshon, 1982; Movshon et al., 1985). This arrangement of direction selective V1 cells providing input to MT cells is motivated by the finding that even though not all V1 complex cells are direction selective, the ones that project to MT typically are (Movshon & Newsome, 1996). The model is able to replicate properties of both, V1 and MT neurons, such as direction and speed tuning and the relation of response to signal strength. Fitting this model to actual data from MT further confirmed that it predicted responses to gratings and plaids (Rust et al., 2006), and even to more naturalistic movies (Nishimoto & Gallant, 2011) very well.

One of the main projection targets of MT is the medial superior temporal area (MST). Neurons in area MST are also motion sensitive and direction selective, but have even larger receptive fields than MT neurons and respond to more complex motion patterns (Tanaka et al., 1986; Saito et al., 1986; Duffy & Wurtz, 1991a, 1991b; Graziano et al., 1994; Mineault et al., 2012). The functional relationship between these neurons' activity and external stimuli is much more difficult to characterize than for V1 or MT neurons, as it cannot be described as a simple tuning for one or two stimulus features. The exact response properties of MST neurons are the main topic of this dissertation and will be covered in detail in the following chapters.

1.5 Overview of chapters

The main part of the dissertation consists of six manuscripts (chapters 2 - 7).

Chapter 2 provides an extensive review of the scientific literature on macaque cortical area MST. We highlight this area's prominent role at the intersection of low-level, bottom-up, sensory processing and high-level, top-down mechanisms that ultimately guide behavior. This review has been published in the *Journal of Neurophysiology* (Wild & Treue, 2021b).

Chapter 3 is a short, more focused review of an article by Sasaki et al. (2017) that discusses how information about self-motion and object motion can be decoded from a population of MSTd neurons with an algorithm that approximates the mathematical process of marginalization. This article has been published in the *Journal of Neuroscience*'s "Journal Club" section (Wild, 2018).

Chapter 4 is a manuscript that accompanies a published dataset of MST neurons' responses to a series of different motion stimuli. This manuscript is under review at *Scientific Data*.

Chapter 5 describes a project that explores the response properties of single MST neurons using a variety of different stimuli and analysis approaches. This project was originally developed by Amr Maamoun and parts of the results have been described previously in his doctoral thesis (Maamoun, 2018). This previous work has been extended to include additional analyses.

Chapter 6 presents preliminary physiological data, simulations, and a proposal for a different approach to characterizing the response functions of MST neurons.

Chapter 7 describes a behavioral experiment with human subjects that aims to explore how different stimulus features affect motion perception.

Chapter 8 summarizes and discusses the findings of the previous chapters and provides an outlook to further questions.

For each chapter a detailed description of the contributions of each person involved in the respective project is provided at the beginning of the chapter.

1.6 Choice of model organism and animal welfare

As most of the research described in this dissertation is based on animal experiments, I would like to briefly discuss the topic of ethics.

Progress in neuroscience – as in all biomedical research – has relied on the use of animal models for centuries. Nevertheless, some people oppose this kind of research on moral grounds. The arguments against animal research typically fall into one of two categories: the *utilitarian* approach argues that the benefits (knowledge and treatment for diseases) need to outweigh the harm (suffering that is inflicted on the animal) in order to justify animal research. Proponents of this approach (e.g., Singer, 1975) argue that animals' capacity for suffering needs to be taken into account when deliberating the ethics of animal research, and typically come to the conclusion that the animals' suffering outweighs the benefits. The *deontological* approach argues that animals are subjects-of-life that have inherent value and therefore must not be used in experiments (T. Regan, 1983).

The critical point in both arguments is whether "human interests should be given greater significance than animal interests" (Brody, 2017), a question that lawmakers in many countries have answered with "yes" (e.g., European Union Directive 2010/63/EU on the Protection of Animals used for Scientific Purposes). This decision does not mean, of course, that animal interests are ignored. The legal framework that regulates animal research is based on the so-called "3R"-principles that had first been formulated by Russell & Burch (1959):

- *Replacement*: the substitution for conscious living higher animals of insentient material;
- *Reduction*: reduction in the numbers of animals used to obtain information of a given amount and precision;
- *Refinement*: any decrease in the incidence or severity of inhumane procedures applied to those animals which still have to be used.

How these principles are interpreted (and ultimately implemented in laws, but also in the laboratory) is in constant flux. One trend in recent years has been to describe "animal research ethics", especially when they concern nonhuman primates (NHPs), as a subfield of research ethics, not only of animal ethics (Arnason, 2020). This shifts questions of autonomy or self-determination into the focus of the debate. Examples of this changed approach to interacting with laboratory NHPs include attempts to assess stress levels

in response to different procedures (Pfefferle et al., 2018) and to provide automatized training protocols with minimal human interaction that provide more freedom and choices to the animals (Calapai et al., 2016; Berger et al., 2018).

The research presented here uses rhesus monkeys for a number of reasons: first, they have a hierarchically structured visual cortex that offers a unique opportunity to investigate how sensory information is processed into more and more abstract internal representations that ultimately guide complex behavior. Second, their brains (especially the visual cortex) share many similarities with the human brain, which allows to draw conclusions that can ultimately be applicable to address neurological and psychiatric diseases. Third, even though the experiments described in chapters 3 and 4 require nothing more of the animal than to keep its gaze on a point on the screen (a task that could potentially also be done by a "lower animal"), they are embedded in a bigger project. The ultimate goal of this project is to explore the effects of cognitive functions, such as attention, on neural activity and therefore requires a model organism that is capable of such functions.

Conscious of the responsibility that is entailed in the privilege of working with NHPs, I believe that the use of rhesus macaques as model organisms for these types of experiments is justified.

Chapter 2

Primate extrastriate cortical area MST: a gateway between sensation and cognition

Benedict Wild & Stefan Treue

Journal of Neurophysiology, Vol. 125 (2021)

As outlined in Chapter 1, the medial superior temporal area (MST) sits at a crucial point in the visual processing hierarchy of the primate brain. This chapter presents an extensive review of the scientific literature on MST.

Acknowledgements

We thank Amr Maamoun for helpful discussion and Bobbie Smith for helpful comments on the manuscript. This publication was supported by the German Research Foundation (DFG) Project 154113120, SFB 889: Cellular Mechanisms of Sensory Processing, project C04: "The pharmacology of attention in fronto-parietal sensory information processing" and DFG Project 211740722, FOR 1847: Physiology of Distributed Computing Underlying Higher Brain Functions in Non-Human Primates, project A1: "Encoding of complex motion patterns in MST of macaque visual cortex and its attentional modulation by the frontal lobe".

Abstract

Primate visual cortex consists of dozens of distinct brain areas, each providing a highly specialized component to the sophisticated task of encoding the incoming sensory information and creating a representation of our visual environment that underlies our perception and action. One such area is the medial superior temporal cortex (MST), a motion-sensitive, direction-selective part of the primate visual cortex. It receives most of its input from the middle temporal (MT) area, but MST cells have larger receptive fields and respond to more complex motion patterns. The finding that MST cells are tuned for optic flow patterns has led to the suggestion that the area plays an important role in the perception of self-motion. This hypothesis has received further support from studies showing that some MST cells also respond selectively to vestibular cues. Furthermore, the area is part of a network that controls the planning and execution of smooth pursuit eye movements and its activity is modulated by cognitive factors, such as attention and working memory. This review of more than 90 studies focuses on providing clarity of the heterogeneous findings on MST in the macaque cortex and its putative homolog in the human cortex.

From this analysis of the unique anatomical and functional position in the hierarchy of areas and processing steps in primate visual cortex, MST emerges as a gateway between perception, cognition, and action planning. Given this pivotal role, this area represents an ideal model system for the transition from sensation to cognition.

2.1 Introduction

Primate cortex consists of well over 100 different areas that can be differentiated on anatomical as well as physiological grounds (Felleman & Van Essen, 1991; Markov et al., 2013, 2014; Paxinos et al., 2000; Van Essen et al., 2012). Around one third of these areas in the human cortex and as much as half of them in the macaque cortex contribute to the processing of visual sensory information (Orban et al., 2004; Orban, 2008). These visual areas are highly interconnected and form a rich network of feedforward and feedback connections between "lower" and "higher' areas. Once visual information coming from the lateral geniculate nucleus (LGN) has arrived in layer 4C of the six-layered primary visual cortex (V1), a hierarchy of brain areas can be determined based on the cortical layers from which projections originate and in which they terminate. Feedforward projections start in superficial layers of the lower area and terminate in layer 4 of the higher area, whereas feedback connections project from deep and superficial areas in the higher area to layers outside of layer 4 in the lower area (Felleman & Van Essen, 1991; Maunsell & Van Essen, 1983a). As one ascends this visual hierarchy, what is represented by the activity within different areas shifts from a representation of low-level features of the 2D retinal image ("sensation") to a high-level interpretation of the multi-dimensional environment and the organisms relationship to it ("cognition") (Treue, 2003).

Areas at the intersection of sensation and cognition are at the heart of a fundamental question of neuroscience: how can the brain extract information about the environment to create an internal representation and subsequently guide behavior? One such area is the medial superior temporal (MST) of the macaque cortex (see section 2.2 for anatomical location and connection to other areas). It is predominantly a visual area that processes information about complex motion patterns, which is typically described in terms of its cells tuning for features of visual motion, such as direction and speed (see section 2.3). But it also uses this information to determine the direction in which the organism is currently moving (section 2.4). Furthermore, it integrates vestibular cues with visual information to improve this representation of self-motion and to tell it apart from the motion of objects in the environment (section 2.5). However, activity in MST neurons reflects not only the integration of sensory input but is also modulated by oculomotor information (section 2.6) and cognitive processes like attention or working memory (section 2.7). Thus, MST is an ideal model system to study the selectivity of individual neurons for complex stimuli, multisensory integration, and sensory-motor transformations and how these processes are modulated by behavioral signals. While this information comes from studies of non-human

primates, there is overwhelming evidence that the human brain contains an anatomically and functionally homolog area (section 2.8).

2.2 Anatomy of the medial superior temporal area (MST)

In the macaque cortex, the MST is located "medial to [the middle temporal area (MT)], along the fundus of the superior temporal sulcus and in places extending several millimeters onto the anterior bank" (Maunsell & Van Essen, 1983a) (Fig. 2.1). An anatomical landmark is a densely myelinated zone (DMZ) that serves to mark the areas border on the upper bank of the superior temporal sulcus (STS) (Boussaoud et al., 1990; Ungerleider & Desimone, 1986). Using anterograde and retrograde tracers, Maunsell and Van Essen (1983a) showed that connections from MT ended primarily in layer IV of MST, which constitutes a feedforward projection (see Fig. 2.2 for a visualization of MST connectivity). The reciprocal connection originated mostly in layers V and VI of MST, which is consistent with a typical feedback projection and is in line with the connectivity patterns of other areas in the hierarchy of extrastriate visual cortex. Boussaoud and colleagues (1990) established many additional cortical connections of MST and the adjacent and highly interconnected fundus of the superior temporal sulcus (FST): MST receives input from hierarchically lower areas V1, V2, V3, the parieto-occipital area (PO), the dorsal prelunate area (DP), and MT. MST has reciprocal connections with area V6 (Galletti et al., 2001), which overlaps with PO (Galletti et al., 2005), and with the dorsal and ventral subdivisions of area V6A (Gamberini et al., 2009; Passarelli et al., 2011). Further reciprocal connections with areas that rank on the same or a similar hierarchical level as MST itself include the ventral and lateral intraparietal areas (VIP and LIP). And it also receives feedback from hierarchically higher areas such as the frontal eye field (FEF; Barbas & Mesulam, 1981) and parts of the inferior parietal lobule (IPL) that have traditionally been referred to as visual area 7a (Andersen, Asanuma, Essick, & Siegel, 1990) and more recently been specified as the "Opt" field and, to a lesser degree, the "PG" field of the IPL (Rozzi et al., 2006). All of these connections are reciprocal: MST sends feedback to the areas from which it receives input and it forwards its output to the areas from which it receives feedback (see Felleman & Van Essen, 1991, Tables 3, 5, and 7 for an overview of all connections, hierarchical constraints, and lists of references).

A study investigating the subcortical connections of MST found reciprocal

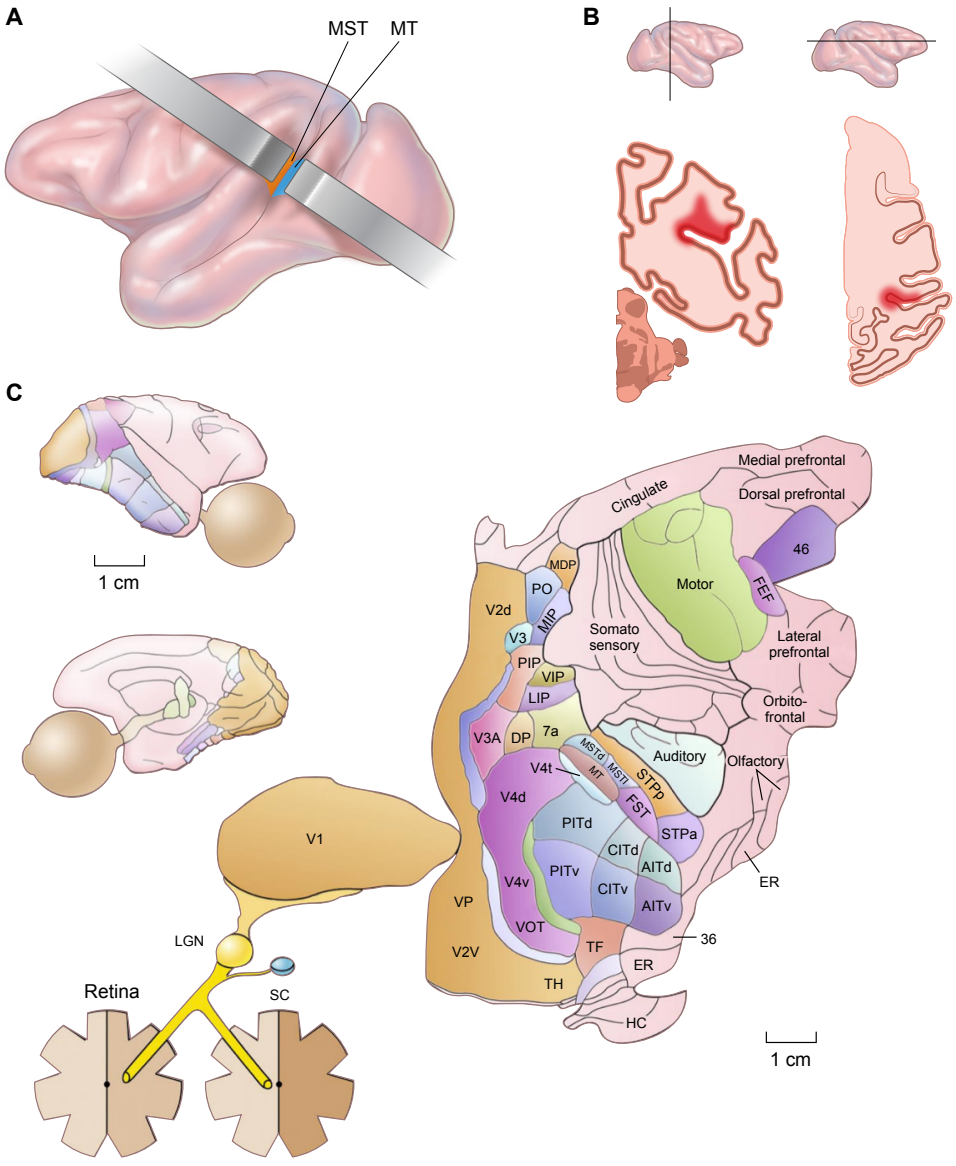

Figure 2.1: Anatomical location of the medial superior temporal cortex (MST) and the middle temporal area (MT) within the superior temporal sulcus (STS). ***A***. MST and MT lie within the STS. ***B***. MST is located on the anterior bank of the sulcus, medial to MT. ***C***. MST is most commonly subdivided into MSTd and MSTl, the exact location of which can be seen in the flattened view of one hemisphere.

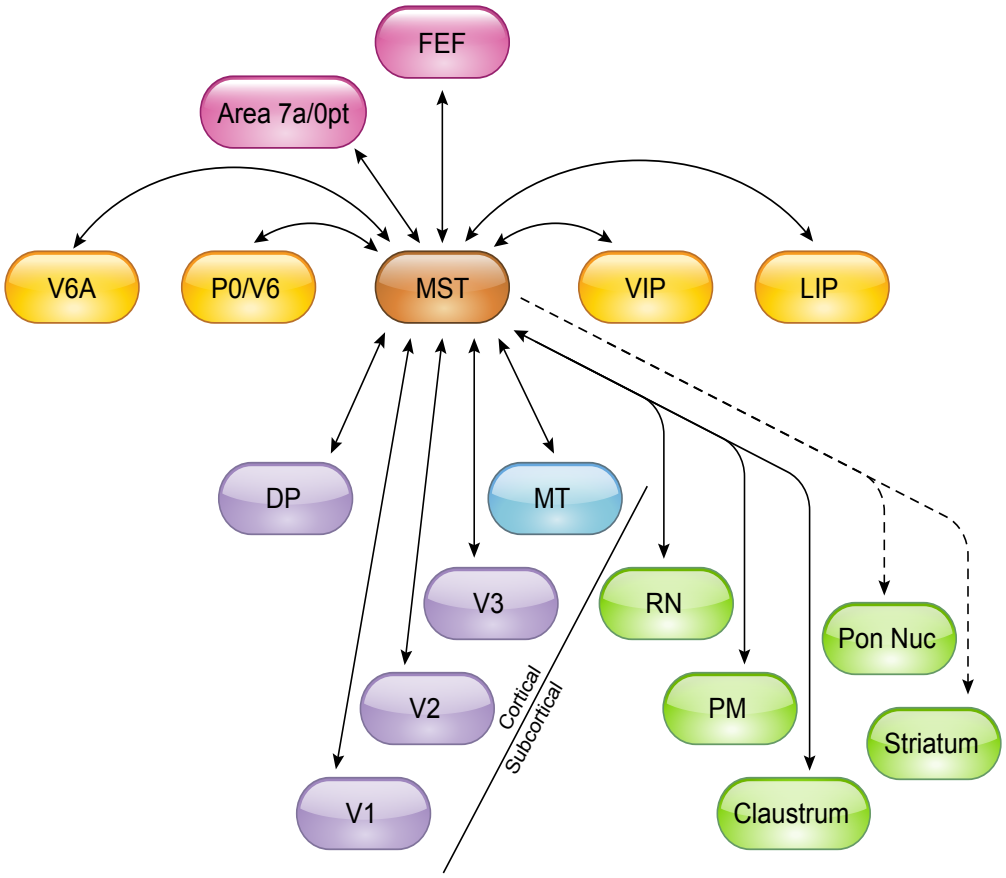

Figure 2.2: Cortical and subcortical connections of MST. The vertical placement of each area is placed with respect at its approximate hierarchical relation with MST. Only connections to and from MST are shown. Dashed lines from MST to Pon. Nuc. and Striatum signify non-reciprocal connections. PO and V6, as well as 7a and Opt, have been put together because of their overlap. Dark-blue indicates cortical areas that are hierarchically lower than MST, light-blue areas are on the same level, red areas are hierarchically higher than MST; green areas are subcortical. FEF: frontal eye field; Opt: cytoarchitectonic field Opt of the inferior parietal lobule; MST: medial superior temporal area; VIP: ventral intraparietal area; LIP: lateral intraparietal area; DP: dorsal prelunate area; PO: parieto-occipital area; MT: middle temporal area; V1/2/3/6: visual area 1/2/3/6; RN: reticular nucleus of the thalamus; PM: medial pulvinar; Pon. Nuc.: pontine nuclei. Cortical connections are primarily based on Felleman & Van Essen (1991), subcortical connections are based on Boussaoud et al. (1992). See main text for details.

connections with the pulvinar, the reticular nucleus of the thalamus, and the claustrum. In addition, Boussaoud et al. found that there were non-reciprocal connections from MST to the striatum and the pontine nuclei (Boussaoud, Desimone, & Ungerleider, 1992). Curiously, the same study reported that injection of anterograde tracers in MST did not show any label in the superior colliculus (SC) which is known to play an important role in eye movements and attention (Gandhi & Katnani, 2011; Krauzlis, Lovejoy, & Zénon, 2013) and is well connected to MT (e.g., Maunsell & Van Essen, 1983a). Apart from the DMZ that marks the areas border on the upper bank of the STS, MSTs boundaries are not sharply defined, which has led to different ways of segmenting the area into subsections. Boussaoud et al. (1990, 1992) differentiate between a more posterior part where cells have receptive fields (RFs) close to the center of the visual field (MSTc) and a more anterior part with cells whose RFs cover the periphery (MSTp). An alternative scheme for dividing MST has been proposed by Wurtz and colleagues, referring to a dorsal-medial subsection on the anterior bank of the STS as MSTd and a lateral-anterior part on the floor of the posterior bank of the STS as MSTl (Duffy & Wurtz, 1991a; Komatsu & Wurtz, 1988b; Newsome, Wurtz, & Komatsu, 1988; Komatsu & Wurtz, 1988a). The counterpart to MSTd has sometimes been labeled the ventral part of MST (MSTv) (Kolster et al., 2009; Pack, Grossberg, & Mingolla, 2001; Tanaka, Sugita, Moriya, & Saito, 1993). Using a variety of different stains, Lewis and Van Essen (2000) differentiate three zones within MST: a dorsal anterior zone (MSTda) that corresponds to what Desimone and Ungerleider (1986) called the DMZ, a dorsal posterior zone (MSTdp) that is located posterior and medial to MT, and a medial zone (MSTm). As of this writing in 2021, a consensus has not yet been reached on how to name the subsections of MST, however the majority of the literature focuses on MSTd. For this review, we will use the naming convention of the respective original publication.

Of the 94 empirical research studies reviewed in the first 6 sections, 69 used rhesus monkeys (Macaca mulatta), 18 were conducted with cynomolgus monkeys (M. fascicularis), 6 with Japanese monkeys (M. fuscata), and 4 with Southern pig-tailed macaques (M. nemestrina) .

2.3 Visual response properties of MST cells

Comparison to MT

Given that MST was originally defined as the main projection target of the middle temporal area (MT), it is not surprising that the two areas share many similarities. MT is a small region located at the posterior bank of the STS and it contains a large number of direction- and disparity-selective neurons that are retinotopically organized, have RFs which are approximately 10 times larger than those of V1 neurons but, like V1 neurons RFs, increase in size with eccentricity (Albright, 1984; Maunsell & Van Essen, 1983c; see Born & Bradley, 2005, for a review). Early studies confirm that MST neurons are also direction-selective, albeit with much larger RFs than MT neurons, often covering substantial parts of the contralateral visual field (Desimone & Ungerleider, 1986; Tanaka et al., 1986). Some respond only to movements of individual luminance bars and not to the movement of wide dot patterns covering a large part of the screen ("figure type cells"). Others show the opposite pattern of responses ("field type cells") or they responded equally well to both types of stimuli (Tanaka et al., 1986). Presumably, figure type cells play a role in detecting the difference between the movements of an object and its background or even in perceiving the boundaries of the object with respect to its environment. Field type cells, on the other hand, which are absent in MT, are involved in the perception of motion of a large part of the visual field, irrespective of individual, smaller objects within that field, such as the motion of the background as one moves through the environment (Tanaka et al., 1986).

Distinct cell types for different motion patterns

However, the most striking difference to MT is that MST contains neurons, which selectively respond to much more complex motion patterns than movement along a straight line. Saito and colleagues (1986) were the first to report three distinct types of cells in MSTd: those that are similar to MT cells and respond preferentially to unidirectional straight movement ("D cells", around 50% of MSTd neurons); cells that respond selectively to radial motion, i.e., an expanding or contracting stimulus ("S cells" for "size change", around 16% of MSTd neurons); and cells that respond selectively to clockwise or counterclockwise rotation in the frontoparallel plane or in depth ("R cells", around 14% of MSTd neurons), leaving 20% unclassified cells. S and R cells are almost exclusively found in the dorsal part of MST whereas the ventral part contains neurons that prefer linear motion of a smaller stimulus and are

presumably more relevant for the perception of object motion (Tanaka et al., 1993). This has led Tanaka and colleagues to suggest that the ventral and dorsal parts are functionally distinct subregions of MST. The selectivity for radial motion in MSTd remains of note since these types of motion patterns are also experienced as one moves through the environment, a phenomenon known as "optic flow" (Gibson, 1950). Importantly, MT cells do not show selectivity for optic flow stimuli (Lagae, Maes, Raiguel, Xiao, & Orban, 1994), suggesting that this property is generated de novo in MST. How MST is involved in the perception of ones own translational movement through the world will be reviewed in detail in the following two sections. For the remainder of this section we focus on general visual response properties, i.e., the neural responses to passively viewed visual stimuli.

Tuning in spiral space instead of distinct cell types

The idea of distinct types of cells (Saito et al., 1986; Tanaka & Saito, 1989) responding selectively to radial, rotational, or translational motion was challenged by Graziano and colleagues (1994). Inspired by the finding that MSTd cells often respond not only to radial, rotational, or translational motion, but to two or all three types (Duffy & Wurtz, 1991a), they hypothesized that these cells might in fact prefer an intermediate form of motion. They defined a continuous circular spiral motion space in which expansion, contraction, clockwise rotation, and counterclockwise rotation can be thought of as the cardinal directions. Intermediate spiral motion patterns can be thought of as a combination of rotational and radial components (e.g., adding a clockwise rotational component to an expanding motion pattern creates an outward clockwise spiral motion pattern, see x-axis in Fig. 2.3 for 8 evenly spaced directions in spiral space). They found that indeed a large majority of neurons had Gaussian tuning curves in spiral space (see Fig. 2.3 for an example), similar to the direction and orientation tuning curves typically found for MT and V1 neurons. A similar study confirmed the selective responses to radial, rotational, and spiral motion and additionally found that almost no MST neurons were selective for deformation (Lagae et al., 1994), thus providing further evidence for a tuning in spiral space. The preferred directions of all tested neurons in Graziano et al. (1994) covered the whole range of directions in spiral space but with a clear bias for stimuli containing an expanding component. Again, this speaks for a role of MSTd in the perception of self-motion, as a forward movement through the environment results in an optic flow pattern that is dominated by an expanding component (explored more in depth in sections 2.4 and 2.5). This tuning is independent of the exact shape of the stimulus, i.e., the preferred direction in spiral space is the

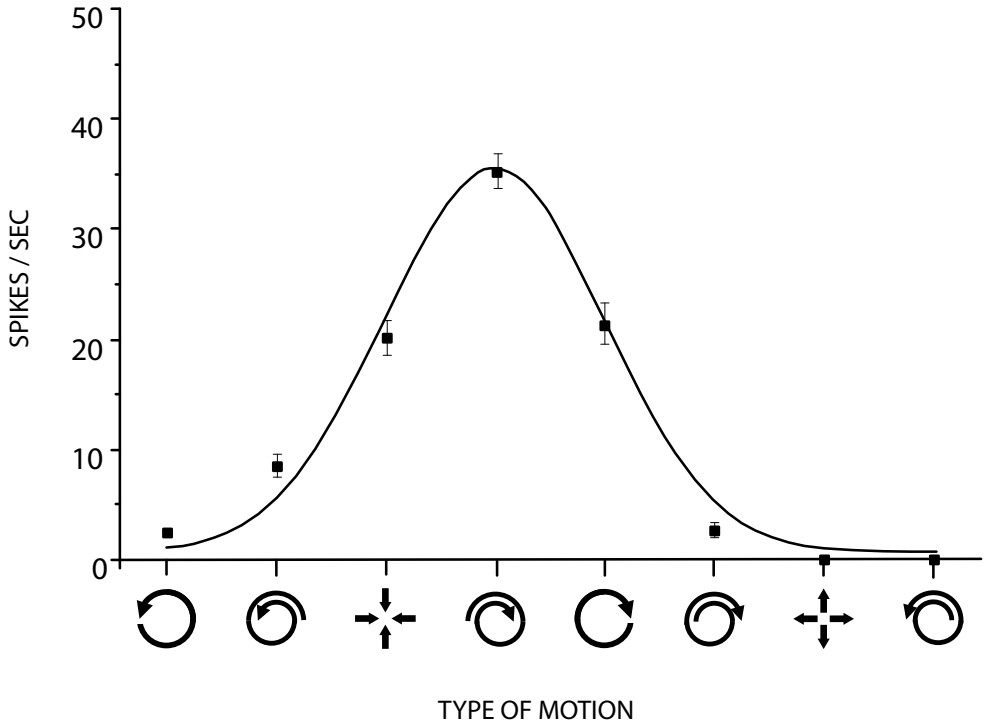

Figure 2.3: Gaussian tuning curve of a cell tuned to a spiral motion that is moving clockwise and contracting simultaneously. The icons on the x-axis indicate the type of stimulus. The error bars show the standard error across 10 trials (Figure reprinted with permission from Graziano et al., 1994).

same for random dot patterns (RDPs) and filled or empty squares (Geesaman & Andersen, 1996). These results suggest that MSTd contains a population of cells tuned to spiral motion directions with their respective preferred directions distributed in spiral space, similar to linear motion preferences in earlier visual areas. There is no evidence for the alternative hypothesis of three distinct subpopulations of cells that decompose complex stimuli into radial, rotational, and translational components. This raises the question whether cells in MSTd are topographically organized according to their preferred direction, similar to the orientation columns in V1 (Hubel & Wiesel, 1962) and direction columns in MT (Albright, 1984). Indeed, both electrophysiological recordings (Britten, 1998; Lagae et al., 1994) and 2-deoxyglucose labeling (Geesaman, Born, Andersen, & Tootell, 1997) indicate that neurons tuned to similar directions in spiral space are clustered in columns in MSTd.

Receptive fields: size, shape, and structure

A recurring theme of this review is that MST neurons generally show more variability and less structure compared to lower areas, such as V1 and MT. In particular, whereas cells in the lateral geniculate nucleus (LGN) and primary visual cortex (V1) are often described as filters that perform relatively simple operations on the visual input (Carandini et al., 2005), this does not hold true for MST, as will be discussed in more detail in the following sections. An antagonistic center-surround structure, which is typically observed in RFs of earlier areas and still present to some degree in MT (Perge, Borghuis, Bours, Lankheet, & van Wezel, 2005; Raiguel et al., 1995), does not seem to be present in MST. Instead, the observation that large stimuli ($\tilde{4}0$ in diameter) are necessary to evoke a MSTd neuron's maximal response suggests that most cells simply spatially sum across their entire RFs (Graziano et al., 1994; Lagae et al., 1994; Tanaka et al., 1993; Tanaka & Saito, 1989). Some studies do report a decrease in firing rate for stimuli that exceed a certain size in some MST cells (Eifuku & Wurtz, 1998; Lagae et al., 1994; Tanaka et al., 1993) and others describe excitatory and inhibitory "zones" of MST neuronal RFs (Duffy & Wurtz, 1991b). Komatsu & Wurtz (1988a) found a reversal in preferred direction in some MST neurons once the stimulus exceeded a critical size, but the preferred direction of a small stimulus did not reverse across different locations within the RF. Instead, the reversal seems to depend on spatial summation over the total RF area that is stimulated. In conclusion, the evidence suggests that any inhibitory mechanisms in MST RFs could provide some form of gain control but do not follow the classical antagonistic center-surround structure. Whereas the size and structure of RFs in early visual areas is well described by simple models, such as a difference-of-Gaussians for LGN cells (Rodieck, 1965), RFs of MST neurons are not only larger, but also more variable in their shape. Fitting receptive fields to a 2-dimensional Gaussian showed MST RFs to be more elliptical or at least less regular than those of MT neurons (Raiguel et al., 1997). Note that forcing RFs into a predefined shape (such as a 2D Gaussian) means that some irregularities in the RF shape get automatically smoothed out. Thus, more elliptical fits could also be a sign that the RF shapes are generally more irregular in MST than in MT.

The relation between RF eccentricity and RF size in the MST is weaker (Tanaka et al., 1986; Desimone & Ungerleider, 1986; Komatsu & Wurtz, 1988b; Luo et al., 2019; Raiguel et al., 1997; Tanaka & Saito, 1989) than that found in the neurons of lower areas (see, for example, Fig. 1 of Freeman & Simoncelli, 2011). Tanaka et al. (1993) even found a negative relation between size and eccentricity for MSTd and a positive one for MSTv. What is generally

agreed upon, however, is that RF sizes in MSTd are much bigger than in MSTl/MSTv (Sasaki et al., 2019). Lastly, electrophysiological recordings of individual cells have not confirmed a well-structured retinotopic organization in MST, as is documented for V1 and MT, in MST. A number of studies report at least crude visual topography (Desimone & Ungerleider, 1986; Komatsu & Wurtz, 1988b; Tanaka, 1983)) while others specifically mention that they found no topography at all (Saito et al., 1986; Tanaka et al., 1986). More recent fMRI studies, however, do provide evidence for a cluster of retinotopic visual field maps in the posterior section of the STS, one of which can be attributed to MSTv, based on its anatomical location (Kolster et al., 2009). Acute single cell recordings and functional imaging both have advantages and disadvantages when it comes to determining the structure of brain regions in visual cortex: the former offers fine-grained information about individual units but samples randomly from the area with high variability between individual recording sessions. The latter can measure the activity across the entire brain within a single recording session, but the spatial resolution is limited as each fMRI voxel represents the blood-oxygen-level-dependent (BOLD) response that is associated with the activity of thousands of neurons, potentially averaging out the variability within this population. The groups of Tsao and Freiwald have had remarkable success in functionally dissecting the inferotemporal cortex by combining fMRI with electrophysiological recordings (e.g., Bao, She, Mcgill, & Tsao, 2020; Freiwald & Tsao, 2010; Schwiedrzik & Freiwald, 2017) and similar methods may be necessary to get a better understanding of the exact structure of MST. Of course, when the receptive fields of single neurons cover as much as a quarter of the visual field, one cannot expect a tessellation of the visual field that is as apparent as it is in earlier areas. At least for those neurons in MSTd with large RFs, the question of whether there is a retinotopic organization or not seems futile. In summary, the size, shape, arrangement, and structure of MST neurons receptive fields cannot be adequately described by simple models or linear relationships. However, recent work looking at more complex models of nonlinear integration of the input that MST receives from MT has shown promising results (Mineault et al., 2012, see below for details) and work along those lines may be helpful in the future.

Speed selectivity

Speed is an integral parameter of motion and how MST neurons respond to different speeds provides important insights about spatial integration of inputs and how complex motion pattern selectivities arise. Specifically, a major question is whether MST neurons simply integrate over local speed

and direction patterns or respond selectively to the overall, global motion patterns inside their RFs. Most cells increase their firing rate with speed until a maximum response is reached and then saturate. A few cells, however, are truly tuned for speed variation; they decrease their firing rate once the stimulus exceeds their preferred speed (A. K. Churchland, Huang, & Lisberger, 2007; Duffy & Wurtz, 1997; Orban, Lagae, Raiguel, Xiao, & Maes, 1995; Tanaka & Saito, 1989). As speed increases, response latency typically decreases in MST and is a bit lower than in MT (Lagae et al., 1994). There is evidence for spatial integration of speed distribution: rotational stimuli normally have a speed gradient, as points on the outer edge of a stimulus need to cover a larger distance to make one full rotation than points close to the center. One study reports that removing this gradient has little effect on the neurons selectivity, suggesting that it is the average speed across the RF, rather than the exact distribution of speeds, that determines its selectivity (Orban, Lagae, et al., 1995). Such a purely spatial speed integration across the RF might be too simplistic, though, as another study did find substantial changes in the responses of up to two thirds of their recorded neurons when removing the gradient (Duffy & Wurtz, 1997). The fact that MT neurons, which provide the main input to MST, are tuned for speed gradients (Martinez-Trujillo et al., 2005; Treue & Andersen, 1996), also makes it likely MST makes use of the additional information about structure that is provided by such gradients. This is one example for the more general question of whether MST neurons analyze motion by parceling complex motion patterns out into smaller, more elementary units, or whether they process them as a unified whole. The contradicting results show that this question has not been fully solved yet and we consider it in more detail below in the context of models of receptive field organization.

Disparity selectivity

Similar to MT (Maunsell & Van Essen, 1983b), a large proportion of neurons in MST is disparity selective (Roy et al., 1992). In other words, their response to a stimulus moving in the preferred direction with the preferred speed (as determined for 0 disparity) varies depending on whether the stimulus is closer (negative disparity, "near cells"), farther (positive disparity, "far cells") or at the same distance (zero disparity) as the fixation point (the "horopter"). Interestingly, Roy and colleagues (1992) report that most cells prefer a non-zero disparity, i.e., only a minority prefers stimuli along the horopter, with the numbers of near and far cells being approximately equal. In around 40% of the investigated cells, the preferred direction switched to the opposite when disparity switched from positive to negative and vice

versa ("disparity-dependent direction-selective" or DDD cells). This feature contributes to determining one's own direction from the motion of stationary objects on the retina caused by the viewers self-motion (Roy & Wurtz, 1990).

Position invariance and receptive field organization

A compelling feature of MSTd cells, first described by Saito et al. (1986), is that of position invariance (Duffy & Wurtz, 1991b; Graziano et al., 1994; Lagae et al., 1994): the observation that a neurons preferred complex motion pattern does not invert, even when local linear motions are inverted (see Fig. 2.4A for an illustration). To test this, Graziano and colleagues (1994) presented stimuli in up to five carefully selected locations within the RF, arranged in an overlapping cloverleaf (to create reversals in local direction) (Fig. 2.4A), to test for preference inversions at these locations. Figure 2.4 shows two cells that preferred clockwise over counterclockwise rotation (Fig. 2.4B) or expansion over contraction (Fig. 2.4C) and retained this preference polarity in all five locations. Graziano and colleagues (1994) found that most responses recorded from MST neurons were position invariant. Using a similar approach, Lagae et al. (1994) found only 40% of their recorded MST neurons to be position invariant; however, they report that nearly all their position invariant cells were located in MSTd, where Graziano and colleagues had also recorded most of their cells, suggesting that position invariance is a dominant (and possibly unique) feature of the dorsal part of MST.

Position invariance is highly informative concerning one if not the most intriguing issue about MST neurons: how can their specific selectivities be generated from their input? The general question of how selectivity for particular features arises has been at the heart of visual neuroscience at least since Hubel and Wiesel (1962) proposed their model of how orientation selectivity in V1 can arise from LGN input. Applying this approach to MT and MST would suggest that several MT cells with properly aligned receptive field locations and preferred directions project onto a single MST neuron which is then selective for that particular arrangement of linear directions (Fig. 2.5). This would be in line with what Duffy & Wurtz (1991b) dubbed the "direction mosaic hypothesis", suggesting that a MST neurons RF consists of properly aligned subfields with translational direction preferences (presumably these subfields would be identical with the RFs of MT neurons projecting onto the MST neuron). And indeed it has been shown that the spatial arrangement of direction components is the most important factor in determining MST selectivity (Tanaka & Saito, 1989). However, this idea is incompatible with the local direction reversals that come with position invariance (Fig. 2.4A): it would mean that for any part of the RF, input needs to be provided by

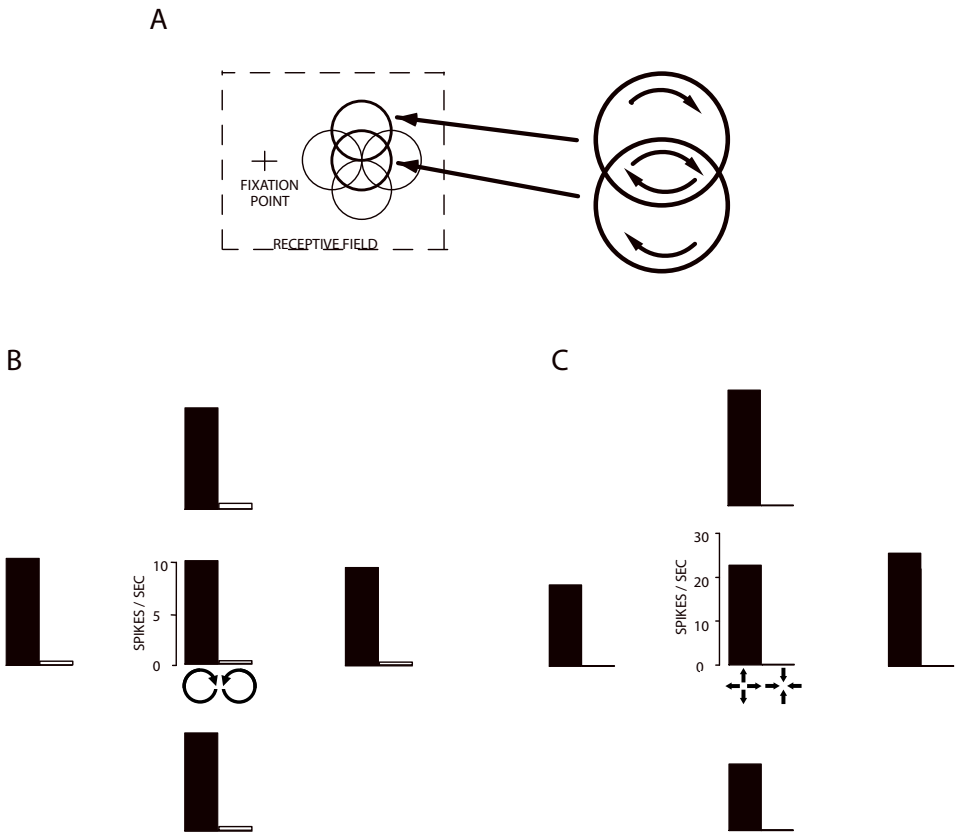

Figure 2.4: Position invariance test used by Graziano and colleagues (1994). ***A***. On each trial, the stimulus appeared at one of five locations within the receptive field (left). As shown in the magnified view (right), in the regions of overlap the direction of motion depended on position. A similar test was used for expanding and contracting stimuli. Black bar = preferred direction, white bar (minimal) = anti-preferred direction. ***B***. Response of a clockwise rotation-tuned cell to clockwise and counterclockwise rotation at the five retinal locations. ***C***. Response of an expansion-tuned cell to expansion and contraction at the five retinal locations. Note that both B and C show responses to the preferred (black bars) and the anti-preferred direction (white bars), but that responses to the anti-preferred direction are close to zero and therefore difficult to see in most plots (reprinted with permission from Graziano et al., 1994).

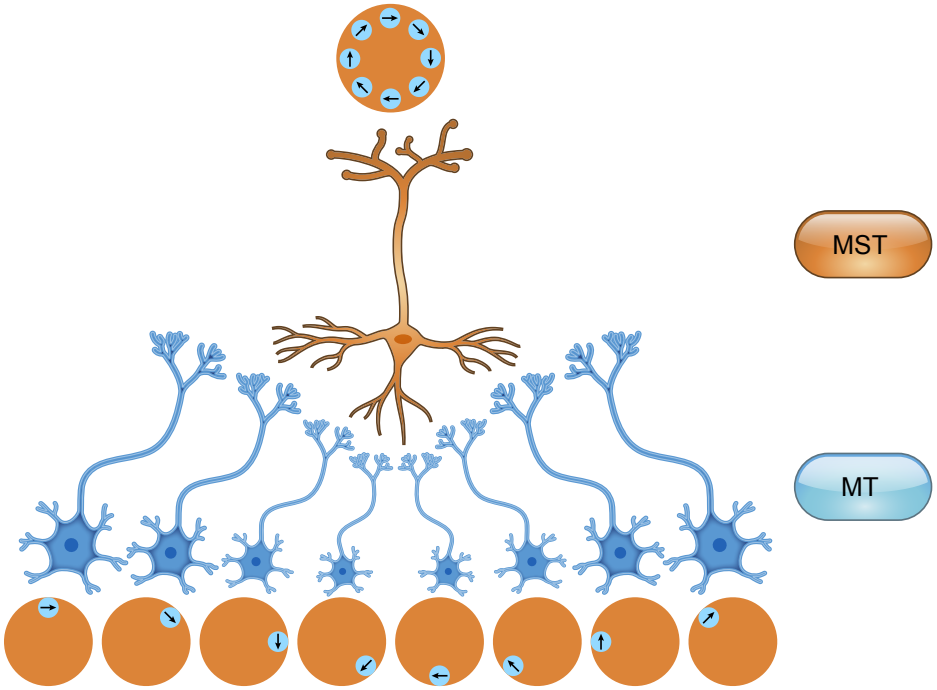

Figure 2.5: A simple architecture to create a clockwise rotation selective receptive field: the receptive fields of MT cells (8 of which are shown here in blue with their preferred directions) are spatially arranged in a way that their preferred directions line up to form a clockwise rotational pattern. If they all project to a single MST neuron (shown in orange) with excitatory synapses, the receptive field of that MST neuron (orange circle) will have a selectivity for clockwise rotation. This model cannot, however, explain position invariant response properties (Fig. 2.4 and text).

multiple MT neurons with different preferred directions. If all of them have excitatory projections to the same MST neuron, the MST neuron would not be selective for one particular direction in spiral space anymore. Several suggestions have been made to address this issue. A "compartment model" divides an MST neuron's RF into overlapping compartments, each of which is constructed from similarly organized MT inputs independently of the other compartments. MT cells whose input creates one such compartment all project onto one branch of a dendrite of the MST cell so that each dendritic branch can be described as a subunit whose activity represents one compartment of the cells RF (Saito et al., 1986).

The alternative "overlapping gradient hypothesis" posits that the RF

consists of excitatory and inhibitory response gradients. The particular arrangement of excitatory and inhibitory gradients with different preferred directions is claimed to account for the selectivities for complex motion patterns (Duffy & Wurtz, 1991b). With a similar idea in mind, Mineault et al. (2012) devised a promising model that can account for the response patterns of a heterogeneous population of MST neurons. They developed a continuous optic flow stimulus that consisted of randomly evolving combinations of translational, spiral, and deformational motion and fitted a neurons response to this pattern to a number of different models, which they then used to predict the neurons response to new stimuli. Not surprisingly, a simple linear receptive field model that compares the stimulus to an internal template was not successful in accounting for the more complex response properties of MST neurons. Instead, a hierarchical model in which an MST neuron linearly integrates the input of several subunits with properties similar to those of MT neurons was able to describe some cells very well, but across the population the predicted responses often deviated substantially from recorded responses. A third model, finally, where the input of the subunits was transformed by a static nonlinear operation with just one free parameter before integration (Fig. 2.6), resulted in remarkably good fits to the data. Such a nonlinear operation could be implemented biologically through inhibitory interactions among MT neurons or synaptic depression between MT and MST. The model found between 2 and 45 subunits for each MST neuron which were mostly excitatory and often had overlapping RFs. It is, of course, likely that each MST neuron receives many more projections from MT than that, but the model convincingly shows an architecture that can explain many of the features of MST neurons.

We would like to reemphasize that the focus of our review lies on studies that include physiological recordings. As is apparent from the examples introduced above, there is also a multitude of pure modeling studies that explore this issue (e.g., Beardsley & Vaina, 1998; Grossberg, Mingolla, & Pack, 1999; Layton & Fajen, 2016; R. Wang, 1995; Zemel & Sejnowski, 1998), which exceed the scope of this review.

Relation of MST activity to perception and behavior

The last question we want to address in this first section is how the activity of MST neurons is related to motion perception on the behavioral level. Two measures based on Signal Detection Theory have been developed to compare behavioral performance in simple discriminations tasks to the responses of single neurons: first, constructing a “neurometric” curve (Britten et al., 1992) allows to compute a discrimination threshold for each neuron, which can be

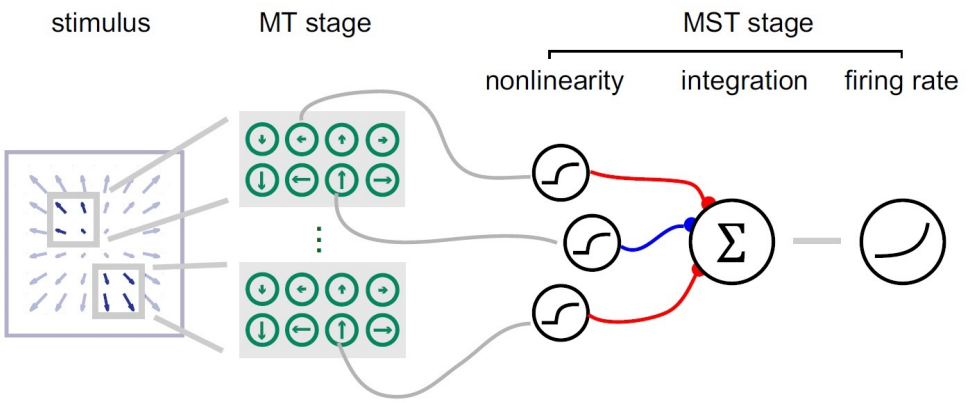

Figure 2.6: Model where a stimulus is processed by groups of MT-like filters. The output of these filters is passed through a nonlinearity and then weighted, summed, and transduced to a firing rate. For each MST cell, the nonlinearity can vary from compressive to expansive and was identical across all subunits (reprinted with permission from Mineault et al., 2012).

compared to the psychophysical threshold of the monkey. Second, choice probability (CP) (Britten et al., 1996; see Crapse & Basso, 2015, for a review, but also Cumming & Nienborg, 2016, and Zaidel et al., 2017, for limitations of CP) is a measure of how the activity of individual neurons is related to the monkeys decision on a trial-by-trial basis. For this measure, one compares the distribution of responses from trials where the monkey chose the neurons preferred feature (here: direction) to the distribution of responses from trials where it chose the anti-preferred feature (here: the opposite or null direction) and calculates the probability that a randomly chosen value from the first distribution is higher than a random value from the second one. Thus, CP values indicate the accuracy with which a neuron's response predicts the monkey's choice, with values around 0.5 representing chance performance and values close to 1 representing nearly perfect prediction accuracy. Both of these measures have been developed to investigate the role of MT in a simple two-alternative forced choice (2AFC) task where the monkey has to report whether a low coherence RDP is moving to the left or to the right (Newsome et al., 1989). This experiment was later repeated to investigate the role of MST in this kind of task (Celebrini & Newsome, 1994). The results were typically quite similar for MT and MST: for most cells, neuronal thresholds were very similar to behavioral thresholds, suggesting that an observer could rely on only a very small group of neurons to make its decision. Similarly,

CP values were significantly above 0.5 in both areas, but did not differ significantly between areas. Furthermore, microstimulation of MST biased a monkeys choice behavior towards the preferred direction of the cells around the stimulation site (Celebrini & Newsome, 1995), again showing a similar pattern of results as MT (Salzman et al., 1992). This suggests that MST does not contribute more or less to this type of simple behavioral task than MT. Interestingly, CP values were not significantly above 0.5 when the monkey had to choose between two opposing directions in spiral space, suggesting that the relationship between spiral motion perception and MST activity is weaker than that between linear motion perception and MST activity (Heuer & Britten, 2004). In contrast, Williams and colleagues (2003) did find a difference in the relation between neural activity and perception between MT, MST, and LIP (lateral intraparietal area): monkeys reported the direction of an apparent motion stimulus, which, on some trials, was constructed so that it could be perceived to move either in a neuron's preferred or its anti-preferred direction. Almost half of LIP neurons and 22% of MST neurons, but no MT neurons showed a difference in activity depending on whether the monkey reported the neurons' preferred or anti-preferred direction. The authors concluded that neuronal activity becomes more aligned with the subjective perception of apparent motion as one ascends through the three hierarchically organized areas. To investigate whether the contributions of MT and MST are necessary for motion perception, Rudolph & Pasternak (1999) lesioned both areas by injecting ibotenic acid. This caused pronounced deficits in motion perception. For stimuli not masked by noise, the monkeys were eventually able to recover some of the impaired perceptual abilities, suggesting that other areas can compensate for the lost function. However, the ability to extract motion signals from noise remained impaired even after extensive training, indicating that MT and MST play an irreplaceable role for challenging motion perception. Lesioning of MT and MST also impaired speed discrimination (Orban, Saunders, & Vandenbussche, 1995). However, a later study showed that very precise lesions of the STS affecting MT, while leaving MST intact, have similar behavioral effects as larger STS lesions that also affect MST. This suggests that the relation of MST activity to behavior is largely inherited from MT, at least for direction discrimination tasks (Lauwers, Saunders, Vogels, Vandenbussche, & Orban, 2000). Yet a different approach to comparing perception with neuronal activity is to examine neural responses during visual illusions where perception diverges from physical stimulus properties. Neural populations, whose activity parallels the illusory perception rather than the physical stimulus, can be considered to be more closely related to behavior than populations whose activity is only related to the stimulus but not to perception. An example of such an illusion

is the "apparent motion" of an RDP whose dots are displaced by a spatial and a temporal separation with successive flashes of each dot. Small temporal separations create "smooth motion", but with increasing temporal (and spatial) separation between flashes (keeping the speed constant), the quality of motion is degraded. Interestingly, observers experience an illusory increase in speed with increasing temporal and spatial separation, but single MT neurons do not parallel this illusion, but rather decrease their firing rates as a function of increasing temporal separation (M. M. Churchland & Lisberger, 2001). However, averaging the responses of a subset of MST neurons allowed to estimate the speed in a way that maps onto the illusion (A. K. Churchland et al., 2007), suggesting that MST activity is closer related to perception in this particular setting than MT activity. In a similar vein, a recent study provided further support for a direct role of MST in motion perception by showing that a subgroup of MSTd neurons respond to illusory rotational or radial motion (Luo et al., 2019). This suggests that MSTd might contribute directly to the perception of these illusions. The exact role that MST plays in motion perception is still not fully understood. The idea that a small, localized set of neurons form the immediate substrate, or "bridge locus" (Teller, 1984), for perception is oversimplified and the neural correlate of any sort of perceptual experience is more likely to be distributed across a number of brain areas (Movshon, 2013). In the case of motion perception, this probably includes areas such as V1, V3, MT, and LIP. Existing detailed models of motion processing typically focus on V1 and MT (Rust et al., 2006; Simoncelli & Heeger, 1998), proposing that MT extracts motion information from V1 and represents velocity in a way that is invariant to other stimulus features, such as spatial frequency or orientation. MST plays an important role within such a multistage model and might perform additional computations that go beyond MT's focus on simple linear motion and extract the corresponding feature-invariant velocity information (Khawaja, Liu, & Pack, 2013). Thus, future work on the neural underpinnings of motion perception should embrace the idea of a network of areas, rather than a unidirectional processing pipeline.

This first section focused on MST with regard to features that are typically discussed in other visual areas, such as receptive field size, tuning for direction and speed, or relation to behavior in simple discrimination tasks. Many of these points will be revisited in later sections. MST cells are not just a slightly more complex step in a one-directional processing pipeline and differ from cells in earlier visual areas in a number of important features, namely

1. highly complex stimulus preferences,

2. no clear retinotopic organization,

3. lack of a suppressive surround structure and

4. questionable relation of receptive field size to eccentricity.

MST seems to be an interface where physical properties of the environment processed by earlier areas are represented in a way that is more directly linked to perception and subsequent action. In that sense, the MST can be thought of as the dorsal pathways analog to the inferior temporal (IT) cortex in the ventral pathway (Graziano et al., 1994). In the next section, we discuss how MSTs unique combination of response properties makes it a central player for the neural processing of self-motion information.

Takeaway:

1. The receptive fields of MST neurons are larger that those of its main input area MT and can cover as much as half of the entire visual field, predominantly on the contralateral side.

2. MSTd neurons are tuned to motion in "spiral space", i.e., they respond preferentially to motion patterns composed of radial and/or rotational directions.

3. Interestingly, this preference for spiral motion is often position invariant, i.e., a neuron's preferred complex motion pattern remains the same in different regions of its receptive field, even if the change in stimulus position causes local motion directions to inverse.

4. How MST neurons integrate the input they receive from MT neurons to create such a position invariant selectivity is not well understood. Evidence so far points to a nonlinear integration of MT responses that are tuned to translational motion, but this is an important area for future research, since it could offer fundamental insights into the neural computation underlying complex stimulus preferences.

5. Elucidating the role that MST plays in motion perception remains a central focus of current research. Understanding how MST works in conjunction with other areas of the dorsal pathway, while taking on a special role as the neural underpinning of at least some motion percepts would help to decipher the role of areas at the interface of sensation and cognition.

2.4 The role of MST in self-motion perception based on optic flow

In both, the ventral and the dorsal visual pathway, the neural representation of the environment gradually shifts from one that (A) primarily reflects low-level stimulus attributes (V1) to (B) a highly specialized representation of selected stimulus features (ventral: V4; dorsal: MT), to (C) a representation that selectively focuses on complex stimulus descriptions of high ecological, social or behavioral relevance (ventral: IT; dorsal: MST). For IT in the ventral stream, prominent examples for species interacting with a complex environment are object recognition (DiCarlo & Cox, 2007; DiCarlo, Zoccolan, & Rust, 2012) and, given the rich social life of primates, facial processing (Freiwald & Tsao, 2010; Kanwisher et al., 1997; Kanwisher & Yovel, 2006; Tsao et al., 2006). What, then, is the equivalent to object representation and face selectivity in the motion domain that leverages the high-level representation in MST? It must be the kind of motion patterns that primates routinely encounter and that require immediate decision making and complex responses. This is the case for optic flow, i.e., the radial patterns that are projected onto the retina during translation of an observer through the environment (Gibson, 1950) and are a major contributor of self-motion perception. Thus, self-motion perception based on optic flow (and separating this self-motion from object motion, which is discussed in chapter 2.5) can be considered MST's core task, in a similar way as object and facial recognition are IT's core tasks. Correspondingly, just like lesions in IT create specific effects on the perception of faces (Barton, Press, Keenan, & O'Connor, 2002; Busigny et al., 2014), and stimulation of face-selective sites biases categorization of noisy images towards a face category (S.-R. Afraz, Kiani, & Esteky, 2006), lesions and stimulation of MST would be expected to impair or bias self-motion perception. Already, the first studies to describe MST neurons selectivity for radial and rotational motion discussed the possibility that these neurons are involved in the processing of optic flow and the analysis of self-motion (Duffy & Wurtz, 1991a; Saito et al., 1986; Tanaka & Saito, 1989). Early psychophysical studies have shown that humans are very good at determining the direction of self-motion from such optic flow patterns, even when the pattern is confounded by eye movements (W. H. Warren, Morris, & Kalish, 1988; W. H. Warren & Hannon, 1988). Since general mechanisms of self-motion perception have been reviewed comprehensively (Britten, 2008; Lappe, Bremmer, & Van Den Berg, 1999), we will focus here on how MST's response properties make it a key area within the brain for solving this problem.

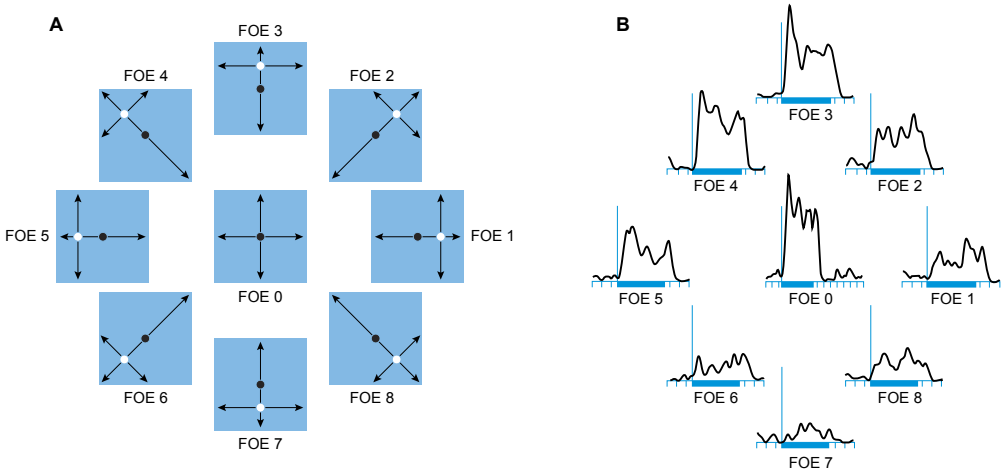

Figure 2.7: ***A***. Nine 76° × 76° visual motion stimuli (box) containing outward radial optic flow (arrows) with FOEs (open spot at origin of arrows) at the fixation point (black dot) or shifted by 30 from the center. ***B***. spike density histograms derived by averaging the responses of a medial superior temporal area (MST) neuron to 6 presentations of the stimulus shown at the corresponding position in A. Vertical bar indicates the 100 spikes/s activity level and the onset of the 1-s visual stimulus the duration of which is marked by the horizontal bar (reprinted with permission from Page & Duffy, 1999).

Tuning for heading direction

To specifically test whether MSTd neurons represent the current heading direction, Duffy & Wurtz (1995) presented monkeys with radial and rotational stimuli that differed in the location of their center of motion. They found that in most neurons the response varied with the location of the center of motion (Fig. 2.7) and that the preferred centers of motion were topographically distributed across the visual field (see also Gu et al., 2006; Lappe et al., 1996; Page & Duffy, 1999; Pekel et al., 1996, for similar findings). This provides strong evidence that the population of MSTd neurons as a whole can encode the position of the center of motion (also called the "focus of expansion", FOE, in the case of expanding stimuli). Thus, presumably, MSTd represents the current direction of heading very well, even if we are not looking where we are heading. Note that variations in firing rate with the location of the center of motion do not contradict the position invariance described in the previous section: position invariance merely states that a neuron's preferred

direction stays the same across different locations *within* the RF. In other words, a position-invariant neuron that prefers expansion to contraction will do so in every part of the RF. The absolute firing rate, however, can still vary across locations and thus allow a neuronal population to represent different centers of motion.

Population coding of heading instead of decomposition of optic flow patterns

It is mathematically possible to decompose any optic flow field into so-called "elementary flow components" (EFC), such as rotation, divergence (i.e., expansion and contraction), or deformation (Koenderink, 1986; Koenderink & van Doorn, 1975). Thus, an early hypothesis was that MSTd performs such a decomposition into EFCs to compute heading direction. However, two findings speak against the decomposition hypothesis: neurons which are selective for one EFC decrease their response when their preferred EFC is mixed with another EFC (e.g., clockwise rotation mixed with expansion from outward clockwise spiral motion) (Orban et al., 1992). Furthermore, as described in the previous section, many MSTd neurons are tuned along a single dimension of spiral motion patterns, rather than representing EFCs (Graziano et al., 1994). If neurons were representing the presence of an EFC, they should respond strongly as long as this EFC is present in the stimulus, even when mixed with other EFCs (Orban et al., 1992). As an alternative, Lappe & Rauschecker (1993a, 1993b) suggest a model consisting of two layers of neurons (such as MT and MST) that can represent heading direction through the population response of the output layer. More specifically, their MST-like output layer represents each possible heading direction with a population of neurons, the summed activity of which provides the likelihood that the respective direction is in fact the current heading direction. This model has received strong support from physiological data (Lappe et al., 1996). In conclusion, the available evidence clearly favors such a population encoding of heading direction, rather than individual MSTd neurons computing heading based on a decomposition.

Effects of microstimulation and inactivation on heading perception

The gold standard for linking neural activity with cognition is to show a causal relationship. Stimulation and inactivation are the methods of choice to document that neural activity is sufficient or necessary for perception or

behavior. To test whether altering the activity of MST neurons is sufficient to modulate heading perception, Britten & Van Wezel (1998, 2002) electrically stimulated the area in monkeys performing a visual heading discrimination task. The monkeys were presented with an optic flow pattern consisting of random dots moving away from a FOE and had to report whether the FOE, which is considered the direction of heading, was to the left or to the right of straight ahead. In a large proportion of experimental sessions, stimulating MST significantly biased the monkeys reports about their heading perceptions, in some cases by more than 5 degrees. Gu et al. (2012) confirmed that microstimulation of MSTd neurons biased behavior in a heading discrimination task and additionally showed that reversible inactivation of MSTd led to strong increases in discrimination thresholds.

Effect of pursuit eye movements on heading representation

So far, we have discussed the highly artificial scenario where a monkey keeps its eyes still by fixating one particular point on the screen for an extended period of time. Only in those cases does the FOE correspond to the direction of self-motion. Eye movements add linear components to the optic flow field, thus shifting the FOE. For example, moving one's eyes to the right (Fig. 2.8E) shifts the retinal image to the left and if this is combined with the expanding optic flow that is associated with forward movement (Fig. 2.8B), it results in an expanding optic flow pattern whose FOE is shifted to the right (Fig. 2.8E, C vs A). Psychophysical experiments have shown that humans can account for these eye movement-induced shifts, but require extraretinal information about eye position to do so (Royden et al., 1992). Since MST also receives and encodes information about eye movements (section 2.6), it is well suited to solve the problem of estimating heading direction from shifted optic flow patterns. It is possible that the tuning of MSTd neurons for heading in fixating monkeys (e.g., Duffy & Wurtz, 1995) could simply represent the position of the FOE on the retina, rather than actual heading. To test this, Bradley and colleagues (1996) observed that monkeys either fixate or perform smooth pursuit eye movements while they were presented with expanding RDPs whose FOE position varied along an axis parallel to the neurons preferred smooth pursuit direction. The authors found different types of cells: "heading cells" showed the same tuning during fixation and pursuit eye movements, whereas "retinal cells" seemed to be responding primarily to the pattern of retinal image motion. Simulating eye movements by adding a linear shift to the RDP while the monkey was fixating made "heading cells" react like "retinal cells":

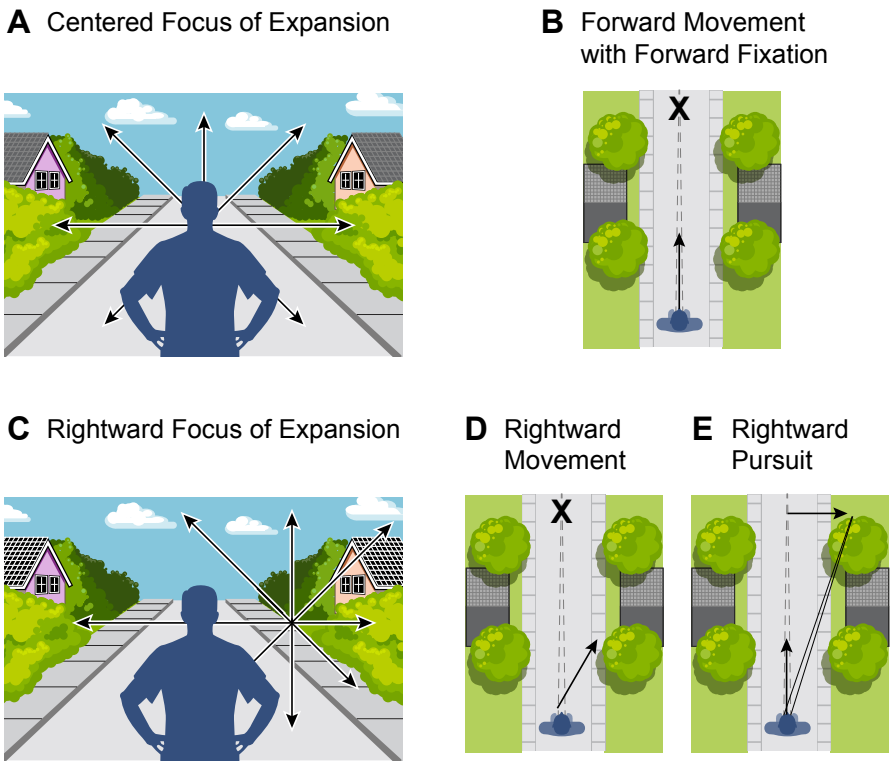

Figure 2.8: Pursuit eye movements shift the retinal image of the focus of expansion (FOE) in the direction of the pursuit, complicating the task of determining heading direction from optic flow. ***A***. Outward radial optic flow with the FOE centered over the fixation point. Directions of motion in the visual field are indicated (→). ***B***. Overhead view of forward self-movement (↑) with forward gaze (—) causing the visual stimulus in A. ***C***. radial optic flow with the FOE shifted to the right. ***D***. Overhead view of right-forward self-movement (↑) with forward gaze (—) that would cause the visual stimulus in C. ***E***. Visual stimulus in C also occurs with forward self-movement (↑) during rightward pursuit (→) (reprinted with permission from Page & Duffy, 1999).

their tuning curves shifted along with the eye movement. This suggests that heading cells have access to information about the eye movement and can use this information to adjust for shifts in the retinal image that are caused by the eye movements. A subsequent study (Page & Duffy, 1999) tested heading tuning during pursuit eye movements in 8 different directions, not just along each neurons preferred pursuit direction. They found that for most neurons, selectivity for a particular heading direction (as simulated by the location of the FOE) was affected by pursuit eye movements (Fig. 2.9). From that, Page and Duffy concluded that individual neurons cannot account for heading detection during eye movements and found instead that a population vector across 196 recorded neurons represents heading well, both during fixation and pursuit (Page & Duffy, 1999). Importantly, the FOE position in head-centered coordinates can be decoded from the population activity even at the single-trial level at an accuracy close to behavioral discrimination thresholds in humans and monkeys (Ben Hamed et al., 2003). This further supports the idea that MST plays a key role in heading perception. However, Bremmer and colleagues (2010) reported slightly different findings: they found that about half of their recorded neurons preserved their selectivity for one heading direction across fixation, simulated eye movements, and real eye movements, similar to Bradley et al.s (1996) heading cells and in contrast to Page and Duffys (1999) findings. But, it should be noted that their measurement of heading selectivity was much more coarse with only three different directions being compared, as opposed to nine different directions in Page and Duffys (1999) study. Furthermore, Bremmer et al. (2010) report that neurons retained their selectivity when the monkey fixated, but the optic flow field was disturbed by adding another flow field that simulated eye movements, which is in contrast to the findings of Bradley et al. (1996) described above. This difference might be explained by the different visual stimuli used in the two studies: whereas Bradley et al. (1996) used classical expanding random dot patterns (white dots moving away from an FOE in all directions), Bremmer et al.'s (2010) stimulus simulated self-motion over a horizontal plane and is more realistic and natural. A recent study attempted to elucidate the relative importance of purely visual, retinal signals and extraretinal efference copy signals for the brains ability to discount the distortions caused by eye movements. Manning & Britten (2019) compared tuning for heading direction in three ways: a normal pursuit, a "simulated pursuit" (during which the monkey fixated, but the stimulus was shifted as if the monkey had made a pursuit eye movement), and a newly developed "stabilized pursuit" (eye movements were compensated by stabilizing the stimulus on the retina based on instantaneous eye velocity). The simulated pursuit condition isolates the effects of retinal signals, as there is no efference

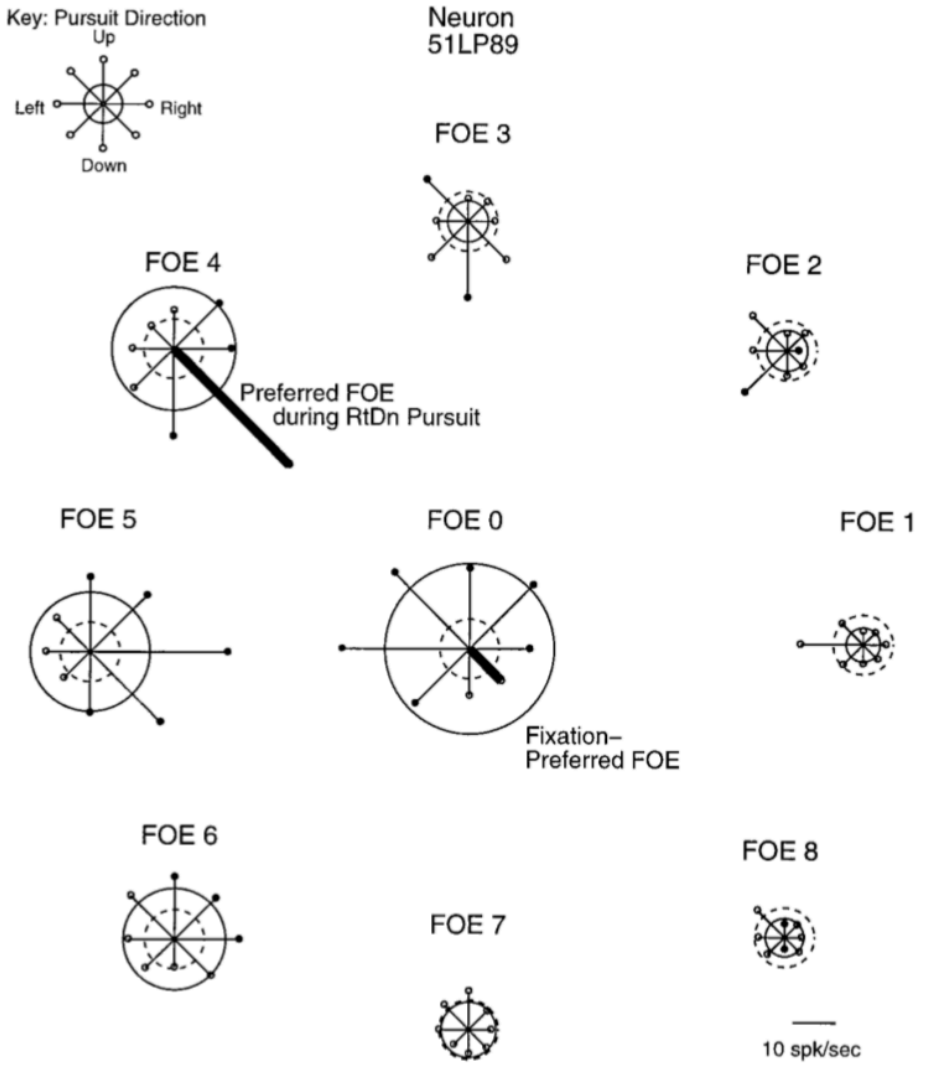

Figure 2.9: Studies of a single MST neuron showing that pursuit shifts which FOE is preferred. Each polar plot shows responses to 8 directions of pursuit (see key at top left) across the FOE located at the position of that polar plot. Solid circles show the response to that FOE during fixation; dashed line circles show the control activity. This neuron preferred the centered FOE in fixation (compare solid circles for each FOE, largest circle is at FOE 0). During right-downward pursuit, this neuron preferred FOE 4 (compare right-down polar limbs for FOE 0 and FOE 4, marked as bold lines, largest is at FOE 4). Vast majority of neurons (93%, 136/146) preferred an FOE, other than the fixation-preferred FOE, during 1 direction of pursuit (reprinted with permission from Page & Duffy, 1999).

copy from the eyes, whereas the stabilized pursuit conditions isolates effects of extraretinal signals, as the eyes are moving but the retinal image does not change. They found that tuning curves shifted very little during stabilized pursuit compared to a fixation condition, which supports the hypothesis that retinal mechanisms alone can explain response stability as these two conditions produce the same retinal image. Furthermore, tuning curves during real and simulated pursuit, which lead to the same, shifted retinal image, were both displaced in a similar manner. All of this suggests that the relative importance of efference copies is rather small compared to the retinal contributions. The differences between these four studies (Bradley et al., 1996; Bremmer et al., 2010; Manning & Britten, 2019; Page & Duffy, 1999) show that heading selectivity during eye movements depends on a complex interaction of the type and direction of the eye movement, the exact stimulus configuration, and the task at hand. There is strong evidence that a population of MSTd neurons can represent the current direction of self-motion even as the visual input is disturbed by eye movements. The exact computational mechanisms that render this possible, as well as their neural implementation, remain an active field of research.

Effect of saccadic eye movements on heading representation

In addition to pursuit eye movements, everyday vision is characterized by ballistic eye movements, so called saccades, which occur multiple times per second during natural behavior. They pose a challenge for heading representation because motion perception is suppressed around the time of saccades (Frost & Niemeier, 2015; Shiori & Cavanagh, 1989). Bremmer and colleagues (2017) showed that a linear decoder that can accurately determine heading direction from a population of MST and VIP neurons, makes systematic errors when analyzing activity during saccades. The decoded heading direction would be compressed towards straight ahead when analyzing the population activity in the time period from just before saccade onset to around 160 ms after. The authors conducted a psychophysical experiment with human subjects who were presented with a short optic flow stimulus and had to perform a heading discrimination task while making an upward saccade. Just like the decoder, the human observers' judgment was biased towards straight ahead when the optic flow stimulus was presented peri-saccadically. This provides strong evidence that the saccade-induced bias in the decoder is not a peculiarity of the decoding approach, but that the information represented in the population activity is truly impaired.

In conclusion, it is clear that MST, together with other, adjacent areas (such as VIP: A. Chen et al., 2011c), plays an essential role in the perception of the direction of self-motion. In fact, this can be considered MST's core task, akin to object recognition in IT, without discounting the possibility that MST may have other core tasks. If it is in fact true, that MST computes and represents ones movement through the environment, it should be able to integrate nonvisual sensory information about self-motion. In the following section, we review evidence that this is indeed the case.

Takeaway:

1. MST's selectivity for "spiral motion" makes it an ideal candidate to process optic flow, the retinal motion patterns during translation.
2. Electrophysiological recordings have confirmed that MST neurons are tuned for heading direction and that this can be decoded from the activity of a population of MST neurons.
3. MST can also represent heading direction when retinal optic flow patterns are distorted by eye movements and the latest results suggest that this representation is based on the retinal input and does not require efference copy information about the eye movements.
4. Elucidating the causal link between neural activity in MST and perception remains a challenge. While the effects of microstimulation and reversible lesioning of MST on heading discrimination support an important role of this area for self-motion perception, the unique contribution of MST to our perception remains to be determined.

2.5 Vestibular tuning and multisensory integration

As outlined in the previous section, the response properties of MST make it an ideal candidate for the neural substrate of self-motion perception. A big question remains: does MST encode purely visual signals about self-motion and passes this information on to downstream areas, which integrate it with information from other modalities to represent heading direction or is it the final, integrated representation of heading direction? If the latter is the case, these neurons should be able to encode heading information based on nonvisual, e.g., vestibular input. It is by no means obvious that

MST might respond to vestibular input, as it had been described as a purely visual area for the first 15 years after its discovery. The literature reviewed in this section show that MST neurons do integrate vestibular information with the visual input. This is further evidence that these neurons reflect an internal representation of the environment that can guide behavior, rather than isolated features of the physical stimulus.

Tuning for vestibular input

Duffy (1998) was the first to test the responses of rhesus monkey MST neurons to optic flow stimuli simulating self-motion and to real motion, both in darkness and in combination with visual stimulation. The real movement was achieved by means of a motorized sled on which the monkeys could be moved in any direction on a horizontal plane. The study confirmed that MST neurons were selective for heading direction in response to visual optic flow stimuli. More importantly, around one quarter of the neurons studied were also selective for the direction of a translational movement in darkness, although the responses were typically smaller in magnitude and less selective than responses to optic flow stimuli. Most surprising, however, was that MST neurons varied widely in how they responded to simultaneously presented visual and vestibular input: Some neurons were strongly tuned for one modality (visual or vestibular) and only responded weakly to the other modality; in that case the response to bimodal stimulation was typically similar to that of the preferred modality alone. In other neurons the response to bimodal stimulation was an additive combination of responses to either modality alone. In a third group of neurons, adding vestibular input to visual input suppressed the response to the optic flow. Notably, the preferred directions for translational movement and optic flow were not related and a substantial proportion of neurons altered their directionality in response to non-congruent bimodal stimulation, i.e., translational movement in one direction and optic flow simulating movement in another direction. The groups of DeAngelis and Angelaki have expanded upon this work, and in a remarkable series of studies, they have explored the integration of visual and vestibular information in MSTd in great detail. Instead of a sled moving on a plane, they have developed a six degrees-of-freedom motion platform which allows for moving a monkey along any arbitrary axis in 3D-space (Fig. 2.10; see Gu et al., 2006, for a detailed description of the setup). Because of the comprehensive reviews of their work (Angelaki, Gu, & DeAngelis, 2009; Fetsch, DeAngelis, & Angelaki, 2010, 2013), we will only briefly summarize their most important results up to 2013 and review newer research published since then in more detail.

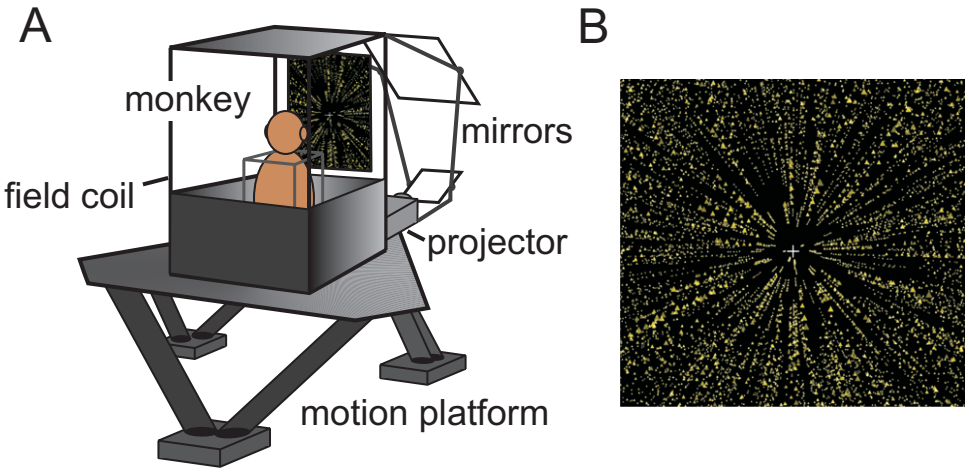

Figure 2.10: Experimental setup used to study visual-vestibular integration by the groups of Angelaki and DeAngelis (e.g., Gu et al. (2006, 2008), Fetsch et al. (2009)). ***A***. The monkey and the display are placed on a six degrees-of-freedom motion platform that can independently provide vestibular (real motion) and visual (optic flow) stimuli. ***B***. Example of an expanding optic flow stimulus (reprinted with permission from Fetsch et al., 2010).

Gu et al. (2006) replicated most of Duffy's (1998) (1998) findings, including the fact that the preferred directions for visual and vestibular stimulation often differed, and found an even higher proportion of MSTd neurons responding to vestibular stimulation (64%, compared to 98% responding to visual stimulation). The reason for this higher proportion of neurons responding to vestibular stimuli is probably the 3D motion, compared to 2D motion on a plane in Duffy's (1998) study. Neurons that were tuned for vestibular stimulation were usually also tuned for visual stimulation so that there were almost no exclusively vestibular neurons and visual responses tended to dominate over vestibular responses in most bimodal neurons (Gu et al., 2006, 2007). Similar to previous findings that MSTd neurons are clustered according to their preferred direction in the visual domain (Britten, 1998; Geesaman et al., 1997; Gu et al., 2012; Lagae et al., 1994), it was found that cells with similar translational or rotational directional preference in the vestibular domain also tended to cluster together (A. Chen, Gu, Takahashi, Angelaki, & DeAngelis, 2008; Gu et al., 2012). Preferred directions in both the visual and the vestibular domain are not uniformly distributed, but bimodal with peaks at 90° to the left and to the right of straight ahead (Gu et al., 2006). Because a neuron is most sensitive at the steepest point of its tuning curve, these

neurons with preferred directions to the left or to the right are hypothesized to be better able to encode heading differences that deviate very little from straight ahead (Gu, Fetsch, Adeyemo, DeAngelis, & Angelaki, 2010).

Multisensory integration on the neuronal and the behavioral level

How is this peculiar set of response properties related to heading behavior? Gu et al. (2007, 2008) trained monkeys to perform a heading discrimination task, similar to the one Britten & Van Wezel (1998) used in their stimulation studies: the animals experienced either real or visually simulated forward motion that had a small rightward or leftward component, or a combination of both (Fig. 2.11), and reported the perceived direction by means of a saccade to the right or to the left. Performance was quantified by constructing psychometric functions (proportion of rightward responses as a function of heading direction) and calculating a discrimination threshold, defined as the deviation from straight ahead that was necessary for the animal to provide reliably correct responses. They found that perceptual thresholds were similar for visual and vestibular stimulation alone (between 1.2° and 4.0°) and improved significantly in the combined condition (Gu et al., 2008).

Neuronal thresholds, determined as described in the first section, were generally worse than the psychophysical threshold for all three conditions (visual, vestibular, combined), suggesting that the animal relies either on the pooled responses of a large population of neurons or gives more weight to the most sensitive neurons (Parker & Newsome, 1998).

To establish a functional link between the activity of MSTd neurons and heading perception, Gu et al. (2007) calculated choice probability (CP) values of MSTd neurons in a heading discrimination task. They found them to be significantly larger than 0.5 in the vestibular-only condition (Gu et al., 2007) and the combined condition (Gu et al., 2008). Additionally, behavioral thresholds for a heading discrimination task predicted from the activity of a population of MSTd neurons were in good agreement with actual psychophysical thresholds that were measured under similar conditions (Gu et al., 2010). Importantly, however, CP values in the combined condition (Gu et al., 2008) depended on the congruency between the preferred visual and vestibular directions of the neurons. Whereas congruent neurons (same preferred direction for visual and vestibular stimulation) had CP values significantly above 0.5, incongruent cells (different preferred directions for visual and vestibular stimulation) had an average CP value slightly below 0.5. This suggests that the monkey relies more heavily on the congruent cells

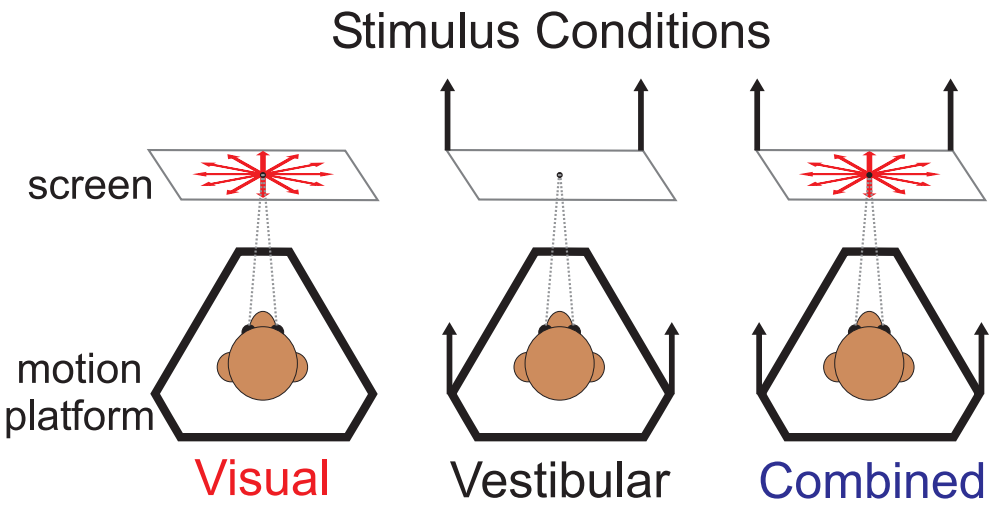

Figure 2.11: Top view of the three stimulus conditions used in the heading discrimination task by the Angelaki and DeAngelis groups (e.g., Gu et al., 2008): visual (optic flow only, indicated by the red expanding optic flow pattern), vestibular (platform motion only, indicated by the black arrows), and combined (optic flow and platform motion). In all conditions, the monkey was required to fixate a central target during the stimulus and then saccade to a rightward or leftward target at the end of each trial to indicate its perceived heading relative to straight forward (one interval version) or relative to the first interval (two interval version). The heading depicted in this schematic is straight forward (0), and thus there would be no correct answer (monkey was rewarded randomly) (reprinted with permission from Fetsch et al., 2009).

during its decision making process and at the same time raises the question of what the purpose of the noncongruent cells is. It should be noted that the interpretation of CP values is complex and a more recent study provided evidence that these values are modulated by both sensory and top-down choice-related signals (Zaidel et al., 2017).

Congruent and opposite cells

The difference between the preferred visual and the preferred vestibular direction in incongruent cells is not random. While Duffy (1998) found no evident relation between the preferred visual and vestibular directions, later studies reported that they tended to be either aligned or pointed in opposite direction, thus classifying cells as “congruent”, “opposite”, and

"unclassified" (Fetsch, Wang, Gu, Deangelis, & Angelaki, 2007; Gu et al., 2006; Sasaki et al., 2017; Takahashi et al., 2007). Furthermore, congruent and opposite cells appear to be arranged in clusters within MSTd (A. Chen et al., 2008). What then could be the purpose of these "opposite cells", considering that their discrimination thresholds get worse and their relevance for behavior (as quantified by choice probability) decreases when visual and vestibular cues are combined, i.e., more information about the environment becomes available (Gu et al., 2008)?

Importance of opposite cells for dissociating self-motion and object motion

It turns out that objects moving through the environment and thereby disrupting the full-field optic flow, can bias heading perception, apparently because the visual systems struggles to dissociate self-motion from object motion. Adding vestibular signals to the optic flow reduces this bias (Dokka, DeAngelis, & Angelaki, 2015). Logan & Duffy (2006) had already shown that an object that disturbs the optic flow pattern (thus suggesting that it's moving independent from the environment) alters responses in MSTd cells that are sensitive to both, optic flow and object motion. In such situations, where visual input is altered and possibly in conflict with vestibular input, opposite cells could help to decompose the overall input into components due to self-motion and components due to object motion. To test this, Sasaki et al. (2017) presented monkeys with the combined visual and vestibular stimulation described above. A cluster of nine spheres, defined by increased dot density (the "object") and moving in one of eight possible directions, was added to the optic flow pattern. They found that indeed, adding a moving object to the optic flow pattern altered the joint tuning for self- and object motion of congruent and opposite cells. In congruent cells, heading tuning was more consistent across different directions of object motion in the bimodal than in the visual condition. In other words, if a cell prefers the same direction for visual and vestibular stimulation, then adding vestibular information can counterbalance the disruption in the visual input caused by the moving object. Tuning for the direction of the moving object, however, was more consistent in the visual than in the bimodal condition, meaning that the addition of vestibular information makes it more difficult for the cell to encode the object's direction. For opposite cells, the pattern reversed: heading tuning was more consistent in the visual than in the bimodal condition, but tuning for the direction of the objects was more consistent in the bimodal than in the visual condition. Importantly, they found that a linear decoder

provides good estimations of heading direction in the presence of object motion (and vice versa) through an approximation of a type of probabilistic inference. This worked only, however, on a population of mixed, i.e., congruent and opposite cells (Kim, Pitkow, Angelaki, & DeAngelis, 2016) and only in MSTd. In MSTl, on the other hand, there are generally fewer cells showing heading selectivity, fewer cells showing vestibular tuning, cells are less able to discriminate directions, and the effects of object motion on self-motion representation are weaker (Sasaki et al., 2019).

Detailed mechanisms of visual-vestibular integration

After establishing that crossmodal information is brought together, a number of studies have investigated how MSTd neurons integrate sensory signals from the two modalities. A number of findings clearly show that visual and vestibular information arrive in MSTd by separate pathways: first, a majority of MSTd neurons are tuned for visual information in an eye-centered reference frame, but in a head-centered reference frame for vestibular information (Fetsch et al., 2007). Second, a bilateral labyrinthectomy eliminates vestibular, but not visual tuning (Gu et al., 2007; Takahashi et al., 2007), which also provides strong evidence that the tuning in absence of visual input is really due to vestibular input and not some unaccounted for additional input (e.g., auditory noise from the motion platform). Third, MT, which provides the major input of visual information to MSTd, does not carry vestibular information (Chowdhury, Takahashi, DeAngelis, & Angelaki, 2009). Instead, Chen et al. (2011a) propose a hierarchical processing pathway for vestibular information, based on the temporal dynamics of direction selectivity, which starts at the parietoinsular vestibular cortex (PIVC), and sends information through the ventral intraparietal area (VIP) to MSTd. But how is the information from these two separate pathways integrated? Morgan et al. (2008) presented monkeys with 64 combinations of visual (8 evenly spaced directions) and vestibular (8 evenly spaced directions) cues. They found that a simple linear combination rule where the response in the bimodal condition was the weighted sum of the corresponding visual and vestibular responses, could account very well for the data. Manipulating the reliability of the visual cue, by changing the coherence of the optic flow RDP, showed that the weights assigned to each modality varied with cue reliability: the visual weight increased and the vestibular weight decreased with increasing visual motion coherence. As mentioned before, monkeys' performance in a heading discrimination task improved substantially in the combined condition compared to visual or vestibular stimulation alone. Similarly, the neuronal threshold (as derived from the neurometric curve) for congruent cells also

decreased in the combined condition, but it increased for opposite cells (Gu et al., 2008). Importantly, both the behavioral and the neuronal threshold in congruent cells in the combined condition were very close to the statistically optimal value. An important prediction of optimal multisensory integration is that different modalities are weighted according to their relative reliabilities. While Morgan et al. (2008) had shown that the relative weights can vary with the coherence of the visual stimulus, this occurred outside of a behavioral task and thus they could not directly test whether the re-weighting is in accordance with optimal integration. Fetsch and colleagues (2009) showed that monkeys performing a heading discrimination task using visual and vestibular cues, are indeed able to weight cues according to reliability in a near-optimal manner. In a follow-up study (Fetsch, Pouget, Deangelis, & Angelaki, 2012) they tested whether a population of MSTd neurons can predict the observed behavioral reweighting and what kind of computations needed to be performed by individual multisensory neurons to account for this. Using maximum likelihood decoding (e.g., Jazayeri & Movshon, 2006), they converted the population response of all recorded neurons into perceptual choices for every trial and calculated psychometric curves based on these calculated responses. The weighting of visual and vestibular cues by this simulated observer matched the monkeys behavior very well if only congruent cells were decoded, suggesting that this subpopulation of MSTd neurons is indeed where the integration and weight adjustment are happening. As in the study by Morgan et al. (2008) the weights of individual neurons varied with coherence and the ratio of vestibular to visual weights was significantly correlated with the statistically optimal weights ratio. More importantly, deviations from optimality matched slight deviations from optimal integration observed on the behavioral level, thus providing further evidence that MSTd is the neural substrate of this integration process. The fact that neurons appeared to be able to change their weights on a trial-by-trial basis indicates that this reweighting does not rely on changes in synaptic weights, as these take place on a slower timescale. A recent model (Ohshiro, Angelaki, & Deangelis, 2011) showed that divisive normalization (Carandini & Heeger, 2012) can account for several empirical findings of multisensory integration, e.g., that multisensory enhancement is stronger for weak stimuli and decreases with increasing stimulus strength ("inverse effectiveness") and that multisensory enhancement works best if cues from different modalities are congruent in space and time. In the model multisensory neurons integrate the responses of two primary neurons, which are sensitive to different modalities but have overlapping response, as a weighted sum. The activity of each multisensory neuron is then divided by the net activity of the pool of multisensory neurons. This normalization can account for the changes in weights assigned to visual

and vestibular cues that were observed when the reliability of the visual cue was altered by changing the coherence of the dot patterns. The authors hypothesize that this is caused by changes in the activity of the normalization pool, which strongly depends on coherence (see Fetsch et al., 2013, for a detailed review). The crossmodal interaction effects predicted by the model have recently been confirmed through recordings of MSTd neurons (Ohshiro, Angelaki, & DeAngelis, 2017). Furthermore, the crossmodal suppression effects predicted by the model were similar in strength for both congruent and opposite cells, indicating that divisive normalization is a general principle of multisensory integration in MSTd and not directly related to a particular behavior, such as heading discrimination or dissociation of self- and object motion. An even more convincing finding in favor of divisive normalization is that unisensory neurons that respond only to visual, but not to vestibular cues show slightly suppressed responses to combined visual and vestibular stimulation (Ohshiro et al., 2017). This can be explained by the model because the normalization pool contains many multisensory neurons and its overall activity is thus influenced by the vestibular stimulation.

This line of research has provided overwhelming evidence that MSTd, which has traditionally been considered a visual area, also processes vestibular information and integrates it with visual cues to represent the current direction of self-motion, or heading. A diverse population of cells that have the same ("congruent") or different ("opposite") preferred directions for visual and vestibular cues allows to optimize heading discrimination on the one hand and to tell self- from object motion on the other hand. The weighting of visual and vestibular cues is flexibly adjustable, depending on cue reliabilities, probably by means of divisive normalization. This line of research has established MSTd as an excellent model system for the study of multisensory integration. Thus, MSTd could, for example, be used to study the neural substrates of age-related changes in multisensory integration, as a recent psychophysical study in humans (Alberts, Selen, & Medendorp, 2019) provided evidence that the weighting of vestibular and visual cues changes with age to compensate for sensory deterioration of the vestibular system.

Note that additional brain areas, other than MSTd, also contribute to heading perception and integration of visual and vestibular information, such as VIP (Bremmer, Klam, Duhamel, Ben Hamed, & Graf, 2002) or the visual posterior Sylvian area (VPS; A. Chen, DeAngelis, & Angelaki, 2011b). We did not discuss these areas extensively, not because we consider them less important, but because our goal here is to review how MST contributes to different functions, such as heading perception, rather than how one specific function is implemented in the brain. The finding that microstimulation and reversible lesioning of MSTd affects behavior in a purely visual task much

more than in a vestibular or a multimodal task (Gu et al., 2012) strongly suggests that other areas can compensate for deficiencies in MSTd. We refer the reader again to Britten (2008) for a comprehensive review about self-motion perception and the different brain areas involved in it.

Role of MST in path integration and spatial cognition

Self-motion and the accompanying change in an organisms position are important aspects of spatial navigation. The neural basis of navigation has been well established in rodents; here, place cells and grid cells in the hippocampus and entorhinal cortex play an important role in the representation of space and ones own position (Moser et al., 2008). Recent studies suggest a similar representation of space and current position by the entorhinal cortex of macaques (Killian et al., 2012) and humans (Jacobs et al., 2013). Ones current position is achieved by previous self-motion. How does the information about self-motion that is encoded by MST neurons contribute to spatial cognition more generally? There is evidence that at least some MSTd neurons can

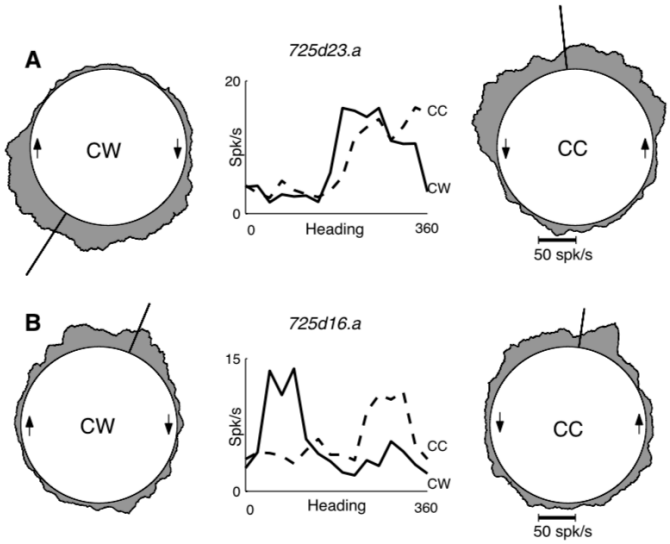

Figure 2.12: Spike-density polar plots (spikes/s) and net vectors (radial lines) with respect to heading. ***A***. Responses of a neuron with similar heading preferences on both paths (clockwise (CW): 237°, counterclockwise (CC): 276°). ***B***. A neuron with opposite heading preferences (CW: 67°, CC: 262°); a place preference for the right front of the room (reprinted with permission from Froehler & Duffy, 2002).

respond differently to the same heading direction depending on the path on which the monkey is moving: when monkeys are moved with a motorized sled on a circular path, so-called "path selective" neurons show larger responses to the same heading when the sled is moved in either clockwise or counter-clockwise directions (Fig. 2.12; Froehler & Duffy, 2002; Page, Sato, Froehler, Vaughn, & Duffy, 2015).

This is particularly interesting, because a previous study found no evidence for temporal integration when measuring responses to a continuously changing optic flow field (Paolini, Distler, Bremmer, Lappe, & Hoffmann, 2000). Either the vestibular input from actual movement is necessary for path integration on the single cell level or path integration only occurs during heading sequences that represent a natural path, which might not have been the case for the artificial setting of a changing optic flow field (Froehler & Duffy, 2002). This suggests that MSTd could be part of a larger navigation and spatial cognition network. Direct connections from the superior temporal gyrus to the entorhinal cortex (Amaral et al., 1983) and a functional MRI study in humans (Wolbers, Wiener, Mallot, & Buchel, 2007) provide additional evidence for such a network.

Takeaway:

1. In addition to visual motion, a subset of MST neurons also responds selectively to vestibular input that provides information about self-motion.

2. Cells that respond to both visual and vestibular information can be "congruent", meaning that they prefer the same heading direction for visual and vestibular cues, or "opposite", meaning that the two preferred heading directions differs by roughly 180°.

3. A population of mixed, i.e., congruent and opposite cells in MSTd appears to play an important role in dissociating self-motion from object motion.

4. The integration of visual and vestibular information has proven to be an excellent model to study multisensory integration and the available evidence supports a flexible weighting of the two modalities, possibly based on divisive normalization, based on cue reliabilities. How the representation of self-motion in MST is combined, with the representation of space and ones own position in the hippocampus and entorhinal cortex, is an important open question.

2.6 Modulation of MST activity by eye movements

In primates and other species with forward-facing eyes and foveal vision systems, perception of motion and eye movements are tightly coupled: When an object that we are looking at moves, we have to move our eyes in order to keep foveating it. How is activity in MST related to these eye movements? Is it modulated by eye movements, and if so: how? What role does it play for the planning or execution of eye movements or both? We already noted that MST neurons representing heading direction can compensate for eye movements, which suggests that they have some information about eye position. The question is whether the contribution of MST to eye movements is mostly perceptual, delivering information about the motion of objects to other areas that compute the actual action plan? Or is MST itself involved in the computation of an action plan?

The first studies of how eye movements are related to the activity of MT and MST neurons occurred in the context of basic motion perception: Newsome et al. (1985) wanted to test whether MT is necessary for the cortical analysis of visual motion. They did this by showing that small, chemical lesions of MT affected a monkeys ability to follow a moving target ("smooth pursuit eye movement") as well as its ability to saccade towards a moving target, but not the ability to saccade to a stationary target. Similar lesioning of MST also suggests that a lesion impairs estimation of the moving stimulus speed and affects targets in the visual hemifield contralateral to the lesioned hemisphere ("retinotopic deficit") or targets moving towards the visual field ipsilateral to the lesioned hemisphere ("directional deficit") (Dürsteler & Wurtz, 1988). These results suggest that the information about the stimulus' motion represented in MT and MST is necessary for the planning of eye movements towards moving targets, but not for computing action plans in general, as saccades to stationary targets were unimpaired. Komatsu & Wurtz (1988b) found that about a third of MT and MST neurons are active during pursuit of a small target in an otherwise dark room so-called "pursuit cells". These pursuit cells also show direction-selective responses to visual motion stimuli during foveation and their RFs typically include the fovea, which is to be expected, since pursuit eye movements will normally start from a point that is currently being foveated. The preferred direction of a moving large-field RDP is opposite to the preferred pursuits direction in a majority of MST pursuit cells (Komatsu & Wurtz, 1988a), which is not surprising, given that as one moves ones eye in one direction (e.g., to the right), the retinal image moves in the opposite direction (in the example, to

the left). To test whether this pursuit signal comes from visual or extraretinal (e.g., corollary discharge or proprioceptive sources) input, Newsome, Wurtz & Komatsu (1988) turned off the pursuit target for a short time interval (the "blink") during the eye movement. While most pursuit cells in MT and some in MSTl show a reduction in firing rate during the blink, suggesting that their pursuit response can be accounted for by visual stimulation from the pursuit target, pursuit cells in MSTd and the remaining cells in MSTl keep their firing rates at a high level during the blink, indicating another, extraretinal source of the signal. In both cases, however, the pursuit-related discharge typically started after pursuit onset, indicating that the activity is not causally involved in pursuit initiation, but rather involved in maintenance of the eye movement. Similar results were reported by Ilg & Thier (Ilg & Thier, 2003) who found that MST, but not MT, pursuit cells responded equally well during pursuit of an "imaginary" target, defined by peripheral cues outside the RF, as during pursuit of a regular target. Microstimulation of MT and MSTl during pursuit movements increases eye velocity when a monkey moves its gaze towards the hemifield that is ipsilateral to the stimulated hemisphere (i.e., leftward pursuit during stimulation of the left hemisphere), but decreases velocity when the eye moves away from the stimulated side (i.e., rightward pursuit during left hemisphere stimulation) (Komatsu & Wurtz, 1989). However, this effect can still be explained by changes in the visual perception of the monkey and does not prove that MT and MST play a causal role in planning or execution of the eye movement. To explore the nature of the pursuit-related signal, Ono & Mustari (2006) recorded MSTd responses during normal pursuit, during pursuit with target blinks (similar to Newsome et al., 1988), and during vestibular ocular reflex (VOR) in complete darkness. As in previous studies, neurons continued their discharge during the blink, but they were not modulated during VOR. The authors interpret this as evidence that the extraretinal signal reflects smooth pursuit or gaze commands, rather than proprioceptive feedback or efference copies, which should also be affected by VOR. However, an alternative explanation that does not assign a motor role to MST could be that efference copies from motor areas are sent to MST only during volitional movements, but not during more automated eye movements, like the VOR.

Sensitivity to eye position

During both fixation and pursuit eye movements, a clear majority of MST single-cell responses to identical stimuli vary with eye position (Bremmer, Ilg, Thiele, Distler, & Hoffmann, 1997). Interestingly, almost all eye position sensitive neurons had their maximum response at eccentric fixation locations.

This could be part of a coordinate transformation computation where information about the environment that has been represented in retinocentric coordinates by earlier visual areas is transformed into different frames of reference (e.g., head-, body-, or space-centered) (Ilg, Schumann, & Thier, 2004). And indeed, Bremmer et al. (1998) have shown that eye position information can be extracted with an optimal linear estimator from the activity of neural populations in different parietal areas, including LIP, 7A and MST. In line with this, Lisberger (2015) suggested that MST plays a similar role for sensorimotor transformation during pursuit eye movements as LIP does for saccades and the medial (MIP) and anterior (AIP) intraparietal cortices do for arm movements. Lisberger's (2015) assessment that the exact role of MST in pursuit remains unknown, still holds true. It is nevertheless clear that MST is the first area along the visual processing pathway that plays a central role for the integration of motion perception and eye movement planning.

Takeaway:

1. The so-called "pursuit cells" in MST are active during smooth pursuit eye movements.
2. In a majority of MST neurons, responses to identical visual stimuli are modulated by eye position.
3. The exact contributions of MST to the integration of motion perception and the planning and execution of eye movements remains to be clarified, potentially offering fundamental insights as to the mechanisms of multimodal integration in higher visual cortex.

2.7 Modulation of MST activity by cognitive processes

The previous section showed that the activity of MST neurons does not solely reflect processing of sensory (visual and vestibular) input in a bottom-up manner, but is also influenced by internal, extraretinal signals, such as eye movements. It is well established that neuronal activity throughout macaque visual cortex is modulated by cognitive factors, such as attention (see Treue, 2001, for a review) and working memory (see Pasternak & Greenlee, 2005, for a review). In this section we review how activity in MST is modulated by attention and (working) memory and how this modulation is similar or different to that of other visual areas.

Attention in visual cortex

Attention is a mechanism that allows the prioritized processing of some sensory stimuli at the cost of impaired processing of other, non-attended stimuli. Visual attention can be either directed towards a certain location ("spatial attention") or towards a specific feature of a stimulus, such as its shape, color, or direction of motion ("feature-based attention"). The neural signatures of visual attention include changes in firing rates (e.g., McAdams & Maunsell, 1999; Treue & Martinez-Trujillo, 1999) and modulated trial-to-trial correlations between neurons (Cohen & Maunsell, 2009; Ruff & Cohen, 2014). Another type of brain activity that has been investigated in the context of attention are local field potentials (LFPs), which can represent the synchronization of neural populations across (low LFP frequencies) or within (high LFP frequencies) brain areas (Buzsáki & Draguhn, 2004). Attention increases synchronization of spikes and LFPs in the gamma frequency band (Fries, Reynolds, Rorie, & Desimone, 2001), changes the phase-amplitude coupling of different LFP frequency bands (Esghaei, Daliri, & Treue, 2015), and decouples spike times from the phase of specific LFP frequency bands (Esghaei, Daliri, & Treue, 2018). There is strong evidence from studies in the ventral pathway that most of these top-down modulatory effects depend on feedback connections (see Squire et al., 2013, for a review) and preliminary evidence suggests that this is also true for area MT in the dorsal pathway (Hüer, 2017; Hüer et al., 2018). MST receives such top-down input from a number of areas in the temporal, parietal and frontal lobes, including the frontal eye field (FEF) (Boussaoud et al., 1990). FEF in particular is considered to be a major source of the modulatory activity that is associated with attention (Buschman & Miller, 2007; Moore & Armstrong, 2003; Moore & Fallah, 2001; Zhou & Desimone, 2011) and its reciprocal connection with MST makes the latter an excellent model to study different effects of attention.

Attention: modulating firing rates

The modulatory effects of attention on individual neurons firing rates increase along the hierarchy of the primate visual system (Maunsell & Cook, 2002; Treue, 2001). As early as in the LGN (O'Connor et al., 2002) and in V1, responses to attended stimuli are higher than those to unattended stimuli (Chalk et al., 2010; Luck et al., 1997; Motter, 1993; Roelfsema et al., 1998). Such gain effects become more prominent along both, the ventral (Cohen & Maunsell, 2011; Fries et al., 2001; McAdams & Maunsell, 1999; Moran & Desimone, 1985; Nandy et al., 2017; Verhoef & Maunsell, 2016) as well as the dorsal visual pathway (Martinez-Trujillo & Treue, 2002, 2004; Mehrpour,

Martinez-Trujillo, & Treue, 2020; Treue & Martinez-Trujillo, 1999; Treue & Maunsell, 1996, 1999).

Given the well-understood sensory characteristics of neurons along the dorsal visual pathway and their midlevel position between early visual cortex and higher association areas, MT and MST in rhesus monkeys have been prime targets for studies assessing the attentional modulation of neuronal responses. In a typical spatial attention paradigm, monkeys are presented with two stimuli, one of which is placed inside the RF of the recorded neuron and the other one at an equal eccentricity, but outside the RF. In a given

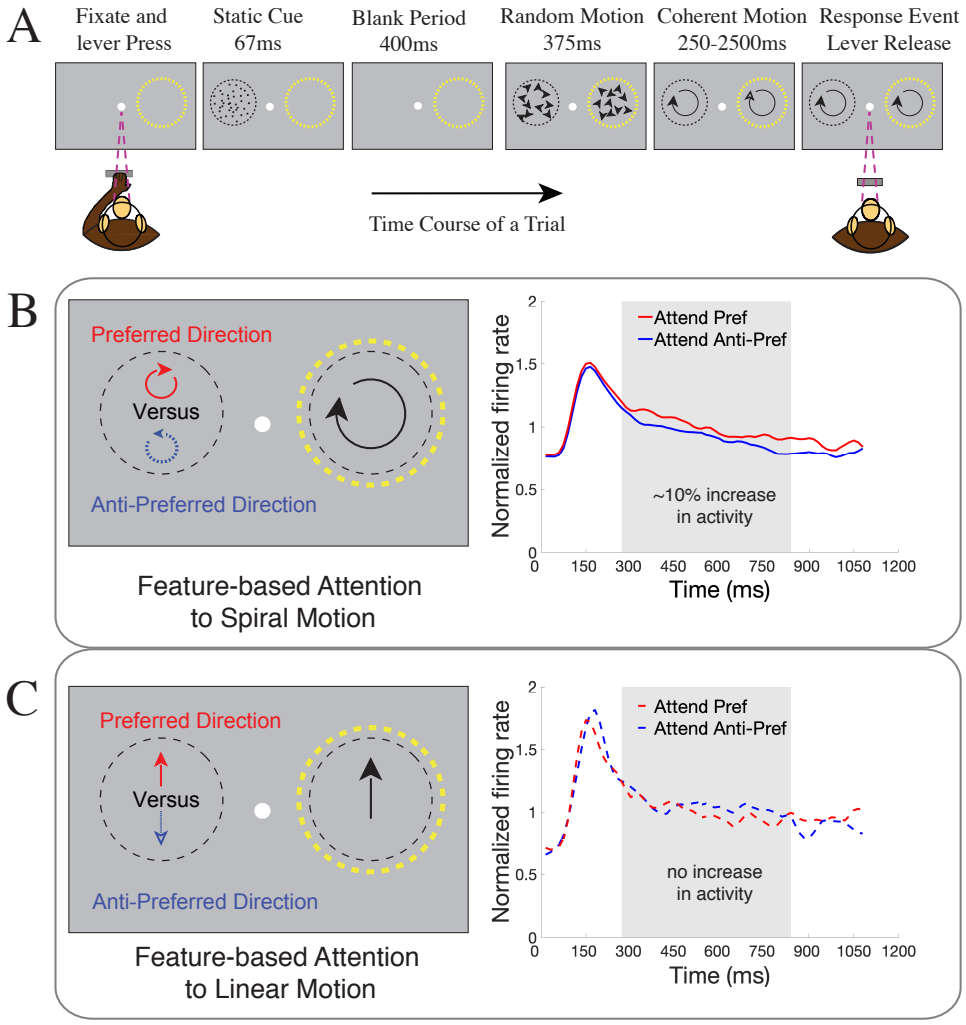

Figure 2.13: *legend on next page*

Figure 2.13: Effect of feature-based attention on an MST neuron. ***A***. Behavioral paradigm: The monkeys initiated each trial by directing and maintaining their gaze on a centrally presented fixation point and holding a touch-sensitive lever. After trial initiation a static cue appeared for 67 ms at an eccentric location, cueing the animal to covertly shift attention towards this target location, either within the receptive field (yellow dotted circle added to figure for illustrative purposes) or in the opposite hemifield. The cue presentation was followed by a blank period for 400 ms to measure baseline activity. To reduce transient motion onset responses, random motion, both inside and outside the receptive field, was presented briefly (375 ms) prior to the onset of coherent motion stimuli (either spiral or linear motion). In order to obtain a liquid reward, the monkeys had to respond to a transient speed increment (250-2500ms after onset) of the target stimulus by releasing the lever, ignoring any speed changes in the distractor. ***B***. Stimulus configuration for the feature-based attention condition with spiral motion. Attention was always directed to the stimulus outside the receptive field (opposite hemifield of yellow dotted circle) to either preferred direction (red) or anti-preferred direction (blue dotted). Inside the receptive field (yellow dotted circle) the stimulus always moved in the preferred direction to ensure a strong sensory response. The right panel shows an example neurons spike density and raster plot for responses while the target stimulus was moving either in the preferred (red) or anti-preferred direction (blue dotted). ***C***. Feature-based attention example for the linear motion configuration. This panel is identical to panel B, except that linear motion stimuli were presented. The right panel shows the neuronal response for the same neuron as shown in panel B, but for linear motion stimuli (adapted from Baloni, 2012).

trial, the animals are cued to attend to one of the two stimuli (the "target", Fig. 2.13A). Importantly, the physical stimulus configuration is identical on trials where the stimulus inside the RF is the target ("attend-in") and those on which the target is the stimulus outside the RF ("attend-out"). Thus, any differences in neural activity represent an internal, attentional signal. Feature-based attention, on the other hand, is typically investigated by presenting a neurons preferred feature (e.g., its preferred direction) inside the RF and cueing the monkey to attend to another stimulus outside the RF that either has the same feature ("attend pref") or the anti-preferred feature ("attend anti-pref") (see left panels of Fig. 2.13B and C for examples of feature-based attention paradigms with spiral and linear motion). To ensure robust sensory responses, attentional studies in MT and MST employ moving stimuli. Across many such studies a multitude of attentional effects

on neuronal responses have been identified and quantified, most notably a carefully orchestrated combination of gain increases for neurons encoding the target stimulus and its features and gain decreases for neurons encoding all other stimuli and features (Malek et al., 2017; Treue & Maunsell, 1996, 1999). Using such an approach, a study (Baloni, 2012) compared the effects of spatial and feature-based attention on linear and spiral motion stimuli for MST neurons tuned to both of these motion patterns (Fig. 2.13). While spatial attention enhanced the responses to these two stimuli in the same way (data not shown), these preliminary results showed that feature-based attention only boosted responses to spiral motion stimuli (Fig. 2.13B), but not to linear motion stimuli (Fig. 2.13C). The authors suggest that their findings provide evidence that the linear motion selectivity observed in MST is not "inherited" from MT (whose responses are affected by feature-based attention to linear motion, Treue & Martinez-Trujillo, 1999) but is generated de novo in MST. The results suggest that spatial and feature-based attention take on complementary roles in MST to combine an unimpeded high gain pass-through processing for sensory information from attended locations in the visual field with an additional feature-based modulation of neuronal responses. This ensures that only those responses to attended features that contribute to perception are boosted by the allocation of attention. This study is the first demonstration of a loss of attentional modulation along the cortical hierarchy and of a selective lack of attentional modulation for just one of the features a given neuron is tuned for. All the studies on attention reviewed so far used highly controlled settings to isolate the effects of attention on elementary sensory processing capabilities of the brain. A different approach was taken by Page and Duffy (2018), who investigated how the neuronal responses to optic flow and real motion are affected when monkeys are engaged in different tasks that require them to focus on either the visual input, or the vestibular input, or a task that is unrelated to the motion stimuli. Monkeys were presented with circular motion by means of either optic flow, or movement of a sled, or both and had to report either a perturbation in the optic flow, or a perturbation in the sled movement, or an unrelated auditory tone via button press. On the single cell level MSTd neurons showed a variety of response patterns that depended on stimulus condition (optic flow alone, sled movement alone, or both) and task (attend to optic flow, attend to sled movement, or attend to auditory tone). On the population level responses were diminished when monkeys attended to the optic flow or the sled movement as compared to when they attended to the tone detection task. While this finding could potentially be explained by the tone detection task being easier, as indicated by faster reaction times, it does suggest that MST responses to heading stimuli are modulated by task

demands. It is an open question: can such a reduction modulation induce systematic biases or misjudgments of self-motion, which would be highly relevant, for example, in the context of driving a car?

For a proper understanding of MST, it is crucial to determine what part of the attentional modulation of activity in MST is inherited from MT, and what part is caused by direct projections from higher areas. Given that there are direct connections between frontal regions, such as FEF and MST, it is plausible to assume that at least part of the enhancement comes directly from a top-down signal to MST itself. To resolve this question, one would have to measure MST responses during an attention task while shutting off the modulation of MT neurons by higher areas. This is difficult to achieve as MT and MST lie next to each other in the cortex and most methods that will affect MT are likely to affect MST as well. However, optogenetics can be used to selectively affect individual cells and some preliminary results suggest that it can be used to selectively shut down the influence of FEF on MT (Hüer et al., 2018).

Attention: reduced burstiness

The attentional increases in firing rates observed in MST are stronger but qualitatively similar to effects found in other visual areas. A recent study found an additional neural signature of attention in MST: spatial but not feature-based attention reduces the occurrence of multiple consecutive action potentials with very short interspike intervals, so-called "bursts" in response to spiral stimuli (Xue et al., 2017). This was particularly surprising as both spatial and feature-based attention increased firing rates in agreement with previous studies (e.g., Treue & Martinez-Trujillo, 1999). Even though the firing rate enhancement was weaker for feature-based than for spatial attention, this difference could not account for the absence of an effect of feature-based attention on the burstiness of MST neurons. Furthermore, the reduction in burstiness could be dissociated from the increase in firing rates, suggesting that they are caused by two different mechanisms. Additional research is necessary to determine to what extent these results are specific for MST or can be generalized to other visual areas. But they strongly suggest that the detailed mechanisms by which attention modulates activity in sensory areas is complex and possibly relies on more than just unidirectional connections from FEF to the respective area.

An open question that, to our knowledge, has not been addressed at all on the neurophysiological level so far is how attention to visual or vestibular cues can modify multisensory integration and heading computation in MST. A psychophysical study (Royden & Hildreth, 1999) showed that allocating

attention across different aspects of the visual input affects object motion perception more than self-motion perception, suggesting that this is a question well worth investigating in more detail.

Working Memory: influence of memory content on sensory responses

The ability to keep a limited amount of information available for a few seconds is essential for any type of goal-directed behavior. It is well established that the prefrontal cortex is an important neural substrate for orchestrating task relevant information and holding stimuli in memory (Miller & Cohen, 2001; Miller et al., 1996; see Constantinidis et al., 2018, and Lundqvist et al., 2018, for a recent debate on the exact mechanisms). A less clear question, however, is to what extent sensory cortical areas are involved in the short term storage of information (see Pasternak & Greenlee, 2005, for a review). An early study recorded the activity of different mid- and high-level visual areas in the ventral and dorsal pathway during a delayed match-to-sample task (Ferrera, Rudolph, & Maunsell, 1994). Monkeys were presented with a sample RDP that was either moving in one of 4 cardinal directions or did not move but had one of four colors. After a short delay period, a sequence of up to four test RDPs, either moving in different directions or with different colors was shown, and the monkey had to respond to the one that matched the sample in either direction or color. In the condition where direction of motion had to be matched, the responses of MT neurons to the four test stimuli were largely independent of the sample stimulus that was kept in memory. In MST and area 7a, in contrast, directional selectivity for the test stimuli became weaker, but activity was more influenced by sample stimulus kept in memory, suggesting that these areas contribute to the maintenance of direction information. This study clearly showed that the memory representation of motion information is by no means trivial: areas outside the PFC appear to be involved, but not necessarily in the same way as they are involved in sensory processing.

Working memory: necessity of MT/MST contributions

Bisley and Pasternak (2000) tried to determine the contributions of MT and MST to working memory by investigating the effects of unilateral lesions on encoding, retention, and retrieval of motion information in a delayed match-to-sample task. They found that monkey performance was impaired, but the exact nature of this impairment depended on the properties of the stimulus and on the task. When presented with a noisy stimulus (a low coherence

RDP) that had to be compared to a coherent stimulus moving in the same or the opposite direction, encoding and retaining was generally impaired by the lesion. This was not the case, however, if the to-be-remembered stimulus contained a strong signal (i.e., a coherent RDP). Comparison of the memorized stimulus with a test was only impaired when the task required the discrimination of similar directions of two coherent RDPs. This suggests that MT/MST is necessary for two tasks: integrating local motion signals across a noisy stimulus for encoding (and possibly retaining) and accurate discrimination of directions. Encoding and retaining of coherent motion for a categorical (left vs. right) task, on the other hand, does not seem to require MT/MST. It is possible that the information can be encoded by direction-selective neurons in earlier visual areas (e.g., V1 and V3) and then be retained as a categorical decision in frontal areas without the involvement of MT/MST. Whether this really is the case could be tested by instructing subjects to maintain either a mental image of the remembered sample or to save the information in an abstract manner (e.g., by verbalizing "up and left"). However, this would be impossible in non-human primates.

Working memory: activity during the delay period

The study by Bisley & Pasternak (2000) was limited in its ability to dissociate the roles of MT and MST in a working memory task as it is very difficult to confine the effects of artificial lesions to one single brain area without affecting surrounding areas. Mendoza-Halliday and colleagues (2014) attempted to find the exact point along the motion processing stream, where direction-selective memory activity emerged by recording simultaneously from areas MT or MST, and the lateral prefrontal cortex (lPFC). Monkeys performed a delayed match-to-sample task similar to the one by Ferrera et al. (1994), where they had to memorize the direction of a sample RDP and, after a short delay report which one of two sequentially presented test RPDs was moving in the same direction (Fig. 2.14A). Neurons were classified as "sensory selective", or as "delay selective" when their firing rates varied as a function of direction during the sample stimulus presentation or the delay period respectively. As expected, nearly all MT and MST neurons as well as 70% of lPFC neurons were sensory selective (i.e., while a sample RDP was shown on the screen, see Fig. 2.14B, D, for example neurons). However, only a third of MST and half of lPFC neurons, but hardly any MT neurons were delay selective (i.e., their firing rate varied with memorized direction in the absence of a visible stimulus, see Fig. 2.14C, E, for example neurons). Furthermore, some MST and lPFC neurons showed strong delay selectivity but weak or no sensory selectivity for direction, suggesting that there is a subpopulation of neurons

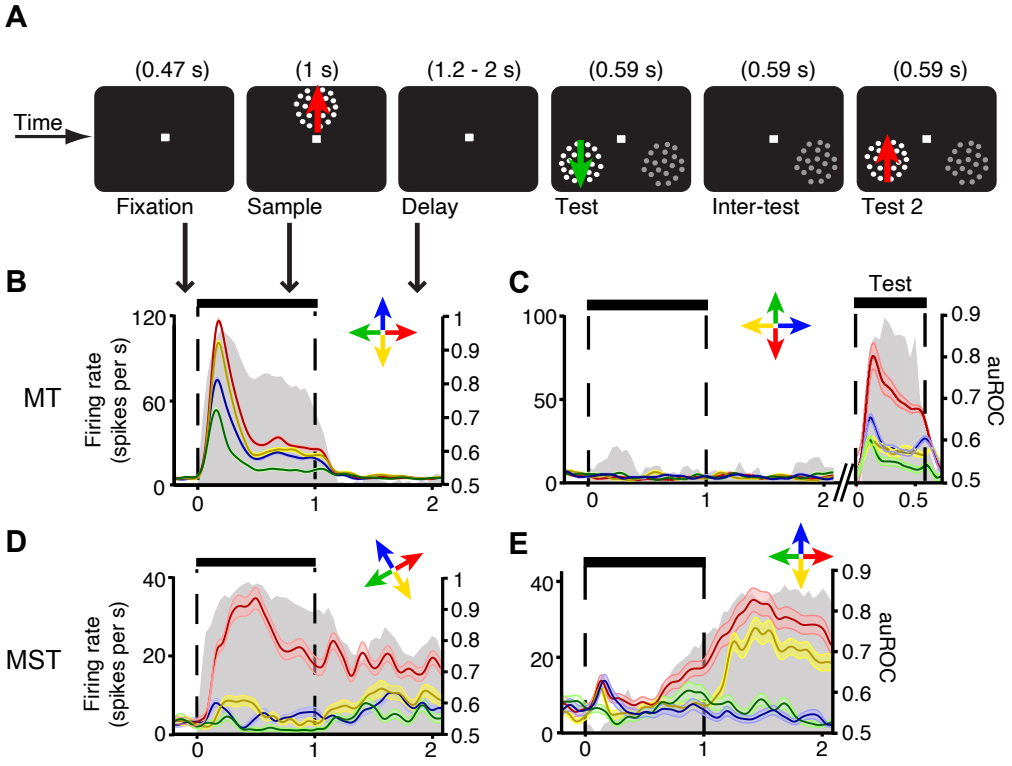

Figure 2.14: Firing rate across task periods for example neurons in MT and MST. ***A.*** Visual display during all task periods. ***B.-E.*** Mean firing rate ($\pm$s.e.m.; $n \geq 50$) over time in trials with each of the four sample directions (color-coded arrows) for neuron examples in MT (B, C) and MST (D, E). Each neuron's preferred direction is shown in red. Gray area shows the corresponding area under the ROC curve (auROC) over time (right axis label). In C, the test stimuli, but not the sample, were placed inside the neuron's receptive field, and colors during the test period represent test directions (reprinted with permission from Mendoza-Halliday et al., 2014).

in these two areas that is primarily concerned with representing a memorized direction rather than the direction of a present stimulus. The delay selectivity in MST and lPFC neurons was also linked to task performance. This result suggests that sustained activity during a short delay period arises quite suddenly between two brain areas that are very strongly connected. Further research will be necessary to investigate how such a pattern can emerge in a presumably small and highly interconnected network. An open question is whether aspects of motion other than direction, such as speed, are also maintained in MST, but not in MT. As was the case for attention, working memory for vestibular information has also been investigated much less than that for visual information and there are, to our knowledge, no studies on the involvement of MST in vestibular memory.

Takeaway:

1. Like in most other areas of the primate visual cortex, firing rates of MST neurons are increased when monkeys direct spatial attention to stimuli inside the neurons' receptive fields.

2. In contrast to the typical effects of spatial attention, feature-based attention was found to only affect neuronal responses to spiral, but not linear motion, suggesting that these two motion preferences play different functional roles.

3. Spatial, but not feature-based attention was found to reduce the occurrence of "bursts" of multiple action potentials with very short interspike intervals.

4. Stimulus-selective activity during the delay period of a memory task appears to be absent in MT, but to emerge in MST, suggesting that this is an important change in the representation of motion from MT to MST.

5. The complex integration of visual information with vestibular signals and their modulation by memory and attentional influences performed in MST underscores this area's potential as an ideal model system and future research focus for the transition from sensation to cognition.

2.8 Human homologs of MST

A main reason why the macaque monkey is such a suitable model organism for cortical information processing is that its brain structure is very similar to that of humans (Orban et al., 2004). Compared to rodent species, non-human primates are more similar to humans in terms of behavior (e.g., the coherence thresholds for motion perception are 2-3 times worse in rats and mice than in primates, Douglas et al., 2006), anatomy (e.g., forward-facing eyes that allow for binocular processing or the multilevel processing pathway for motion processing V1-MT-MST that has no equivalent in the rodent brain) and physiology (e.g., a larger part of motion processing occurring in the retina and V1 in rodents, as compared to primates, Marques et al., 2018). Findings about the non-human primate brain are therefore thought to be more directly relevant for an understanding of the human brain (Buffalo, Movshon, & Wurtz, 2019; Mitchell et al., 2018; Roelfsema & Treue, 2014). In this section, we review the evidence that the human cortex contains a homolog to macaque cortical MST. The interspecies similarities that we describe below underscore the relevance of understanding macaque MST for our understanding of human vision and cognition.

Psychophysics

Visual psychophysics in healthy human subjects can be used to determine processing "channels" (Campbell & Robson, 1968; Lee, 2011) or "detectors" (Ringach, 1998) for specific visual features, such as orientation, luminance, or motion direction. While it is not always straightforward to exactly map such channels onto neural structures, they clearly suggest specialized modules in the human brain that underlie the processing of these features. The existence of such channels can be demonstrated by showing that performance in visual detection or discrimination tasks depends critically on individual features of a stimulus. For example, a classic study showing that adaptation depends on the spatial frequency of a stimulus suggests that there are processing channels in the human visual system that are selective for spatial frequency (Blakemore & Campbell, 1969). In a similar way, Regan and Beverley (1978) provided evidence for "looming detectors", i.e., processing channels in the visual system that selectively respond to changes in size (which can be described as expansion and contraction), separate from motion information. It was not clear, however, that their adaptation paradigm probes a putative human homolog of MST, rather than earlier areas, such as V1 or

MT[1] (Burr & Thompson, 2011). Therefore, Morrone and colleagues (1995) tested the integration across local motion signals by measuring coherence thresholds in a direction discrimination task. They showed that motion sensitivity increases with the area of a random dot pattern, suggesting a processing channel that integrates signals across the visual field. Importantly, for circular or radial motion, such a spatial summation cannot be explained by channels that are tuned to local linear motion, as different subsections contain motion in orthogonal or even opposing directions. Instead, there must be a neural mechanism that integrates different motion signals across sectors (a putative MST-like channel). Furthermore, sensitivity is lower when sectors of the RDP that contain no signal dots are filled with noise dots as compared to when they are empty, providing additional evidence that this integrating mechanism sums signals across the entire stimulus. In a second experiment, contrast was varied for the entire stimulus while coherence was held constant. Only a small effect of stimulus area on contrast sensitivity was observed, suggesting that no summation takes place and sensitivity is limited by an earlier stage with a contrast threshold (e.g., V1). A follow-up study (Burr, Concetta Morrone, & Vaina, 1998) using annuli confirmed the summation across large regions of space (as far as 72) and provided evidence that the integrator mechanism relies on neurons with large receptive fields. Based on these results, Morrone, Burr, and colleagues suggest a two stage process of complex motion processing in the human brain: The first stage is a number of independent local motion detectors (e.g., V1 or MT cells) that respond to motion in a small part of the visual field and that limit overall stimulus visibility through a contrast threshold. The second stage is an integrator mechanism over the local motion detectors (presumably a human homolog of MST) that is able to analyze more complex motion patterns across an extended region (see also Vaina, 1998, for a review on the physiology and psychophysics of complex motion perception). Further similarities between the response properties of MST cells in the macaque cortex and an optic flow channel in the human visual system were demonstrated by Snowden and Milne (1996). They showed that adapting to a spiral motion RDP (as described in section 1) elicited aftereffects that were selective for the adapting direction and position invariant, which agrees well with the properties of MST neurons described by Graziano and colleagues (1994). We have highlighted a few selected psychophysics experiments that specifically aimed at providing evidence for an MST-like processing channel in the human brain. To review

[1]It should be noted, in all fairness, that when the study by Regan & Beverley that Burr & Thompson refer to was published, MST had not been described yet, MT had not been explored in great detail yet and even the concept of a dorsal pathway had not yet been proposed.

the entire field of motion psychophysics would clearly exceed the scope of this review and we refer the reader to existing reviews covering this field in more detail (Burr & Thompson, 2011; Nishida, 2011; Nishida, Kawabe, Sawayama, & Fukiage, 2018).

Functional Imaging

Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) provide a more direct measure of motion perception in the human cortex (always keeping in mind that while BOLD activity is highly correlated with neural activity (Logothetis et al., 2001), the two are not identical). A motion-sensitive area in the inferior temporal sulcus (ITS) of the human cortex, considered a homolog to macaque cortical areas specialized for motion processing, has been established quite early (Tootell et al., 1995). This complex is often referred to as hMT/V5 ("human MT", e.g., Orban et al., 1998), MT+ or hMT+ ("human MT" with the "+" suggesting additional areas being included; e.g., Beauchamp, Cox, & DeYoe, 1997), or simply "MT-MST" (e.g., O'Craven, Rosen, Kwong, Treisman, & Savoy, 1997). All of these names acknowledge that this complex probably contains multiple areas and a number of different fMRI studies strive to differentiate these areas, each leveraging one specific difference between MT and MST. Morrone et al. (2000) made use of the fact that MST, but not MT neurons respond selectively to circular and radial motion trajectories. They found a part of the V5/MT complex along the sulcus that separates Broadmanns area 19 (V3, V4, V5) and 37 (fusiform gyrus) that showed activation when contrasting responses to circular, radial, and spiral RDPs with responses to randomly moving noise RDPs. This area was distinct and on average more than 1 cm removed from another area in the V5/MT complex that was selectively activated by contrasting translational motion with noise RDPs. Dukelow and colleagues (2001) made use of the fact that MST, but not MT receptive fields extend into the ipsilateral visual field and MST, but not MT neurons receive extraretinal smooth pursuit eye movement signals (as reviewed in section 4). They found that ipsilateral optic flow stimuli produced activation at the anterior end of the MT+ complex (putative MST) while the posterior part of the MT+ complex (putative MT) was only activated by contralaterally presented stimuli. Nonvisual smooth pursuit eye movements in darkness activated a small volume in the anterolateral section of the MT+ complex that responded only weakly to contralateral or ipsilateral motion. The authors thus conclude that the human MT+ complex can be divided in three areas:

1. an anterior part that responds to contra- and ipsilateral optic flow and

can be considered a homolog of macaque MSTd,

2. an anterolateral part, slightly inferior to the first, that is selectively activated during nonvisual pursuit which shares similarities with macaque MSTl, and

3. a posterior part that only responds to contralateral motion stimuli and can be considered a homolog of macaque MT.

Peuskens and colleagues (2001) investigated brain regions involved in heading estimation by presenting observers with ground plane optic flow patterns. Contrasting activity in a heading task with a control condition, they found that the heading task activated hMT/V5 as well as a more ventrally located region, which they called the "inferior satellite of hMT/V5". They suggest this region to be a likely candidate for the human homolog to macaque MSTd. Huk et al. (2002) made use of the fact that MT has a much more fine-grained retinotopic map than MST and that MST receptive fields are much larger than MT receptive fields and often extend into the ipsilateral visual field. They assessed retinotopy with a motion defined wedge that rotated through the visual field and compared ipsi- and contralateral responses to stimuli restricted to one hemifield. A subregion of the hMT+ complex on the posterior (or ventral) bank of the ITS showed response modulations to the rotating wedge consistent with a retinotopic organization (presumably MT) and a separate subregion on the anterior (or dorsal) bank of the ITS showed strong activation in response to to ipsilateral, peripheral stimuli (presumably MST) (Fig. 2.15). This division of the hMT+ complex in two areas based on retinotopic organization, responses to ipsilateral stimuli, and responses to optic flow stimuli was also confirmed by Smith and colleagues (2006).

To differentiate between brain areas that are selective for optic flow from those that respond to complex motion more generally, Wall & Smith (2008) compared responses to a single, large patch of optic flow or an array of nine similar patches that did not indicate egomotion. They found that MST responded well to both types of stimuli, but significantly better to the egomotion-consistent, single optic flow stimulus. Similar findings in the macaque brain, using the same stimuli (2017), provide further evidence for the similarities between human and macaque MST. More recent approaches in functional imaging have moved beyond trying to isolate individual areas and instead focus on networks of areas. A retinotopic mapping in combination with stimuli designed to test motion and shape sensitivity identified 18 retinotopic occipital regions, including 4 regions that constitute the human MT/V5 complex (Kolster, Peeters, & Orban, 2010). Comparison with similar fMRI studies in macaques (Kolster et al., 2009) suggests that one of these four

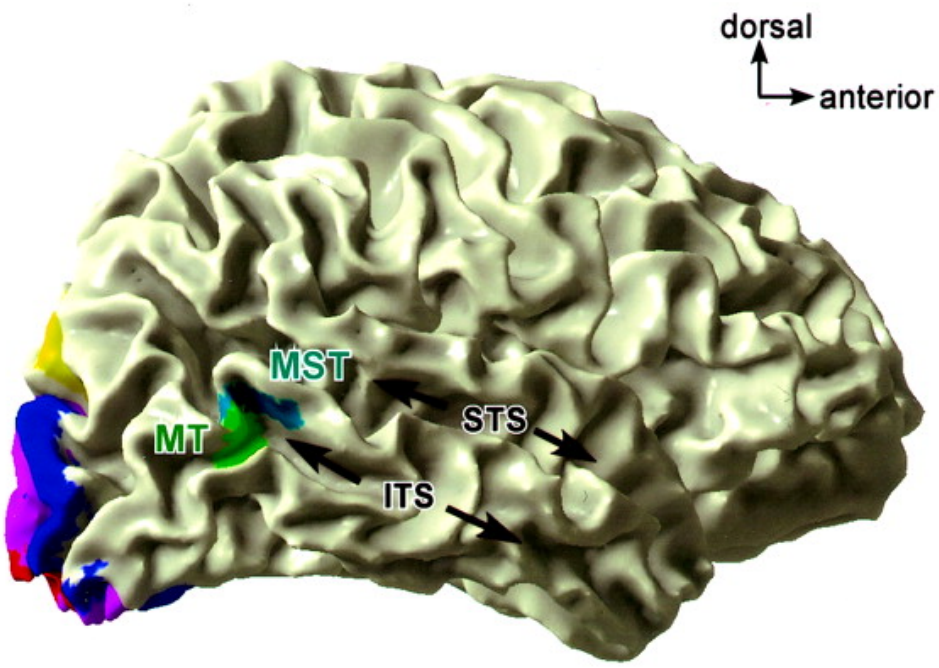

Figure 2.15: Position of MT and MST, viewed on a 3D cortical reconstruction, for one subject (left hemisphere). MT (green) falls on the posterior bank of the occipital continuation of the ITS, whereas MST (turquoise) falls on the anterior bank. The STS is indicated for reference. Other visual areas are shown for reference: V1, red; V2, purple; V3, blue; V3a/b, yellow (reprinted with permission from Huk et al., 2002).

regions of the MT/V5 complex is a putative homolog of MSTv (but not MSTd) as they share similar topographic organization and topological neighborhood and in accordance with other studies, this putative MSTv is located anterior to MT/V5. In conclusion, while the overall location of the hMT+ complex differs slightly from the location of the MT/MST complex in the macaque cortex, fMRI studies provide overwhelming evidence for a posterior and an anterior area within the hMT+ complex that show strong anatomical and functional similarities with macaque MT and MST, respectively.

Transcranial magnetic stimulation

Strong et al. (2017) selectively disrupted neural activity in human MT or MST using transcranial magnetic stimulation (TMS) to test how these two

areas differ in their contribution to the perception of different types of motion. Participants had to discriminate the translational (up vs. down), radial (expansion vs. contraction) or rotational (clockwise vs. counterclockwise) direction of a low coherent RDP presented in the periphery while MT or MST in the hemisphere contralateral to the stimulus were stimulated with 5 TMS pulses during stimulus duration. Stimulation of either area impaired performance for the translational direction discrimination, but only stimulation of MST impaired performance in the radial and the rotational direction discrimination tasks. The authors argue that this indicates a serial processing stream where information is passed from MT on to MST, which then integrates this information to represent more complex forms of motion. This reflects a broad consensus on how the larger receptive fields in MST are built from a mosaic of receptive fields, representing individual MT neuron input. On the other hand, studies like the one cited above (Baloni, 2012) suggest that neuronal tuning aspects in MST could also be generated from V1, V2, and V3 inputs, bypassing MT. This is supported by the lack of impairments in perceptual performance for radial and rotational movements when MT is interrupted by TMS.

Functional similarities between human fMRI and monkey physiology in studies of higher cognition

Additional studies suggest that some of the MST properties that have been described in the previous sections for the macaque apply to the human as well: Using fMRI, OCraven et al. (1997) showed increased activity in MT/MST when subjects attended to moving dots compared to when they attended to simultaneously presented stationary dots. This suggests that attention modulates MT/MST activity in humans in a similar manner as it does in macaques (Treue & Maunsell, 1996). Thus, a human homolog of macaque cortical MST that is distinct from a homolog of MT has been established by making use of a number of ways in which the two areas differ, such as receptive field size and responses to linear vs. spiral motion. A third example for a categorical difference between MT and MST is the activity during the memory period of a delayed match-to-sample task that was described in the previous section (Mendoza-Halliday et al., 2014). It is an interesting open question whether this finding can be replicated in humans as well. Answering the question requires a high spatial resolution to distinguish between MT and MST, but does not rely on a millisecond temporal resolution, as the delay period can last for several seconds. Therefore, fMRI would be a suitable method to investigate this question.

In conclusion, there is overwhelming evidence that macaque cortical MST has a homolog in the human cortex. Thus, MST is not only suitable as a model area for studying a variety of general sensory and cognitive mechanisms, but also allows us to draw strong conclusions about neural mechanisms underlying human vision.

Takeaway:

1. Psychophysical experiments suggest that the human visual system contains specialized processing "channels" or "detectors" for rotational and radial motion patterns that share similarities with the processing in macaque MST neurons.

2. Functional MRI studies have identified areas within the human MT+ complex that share similarities with macaque MST, such as selective responses to optic flow and large receptive fields and therefore constitute a likely homolog.

3. Transcranial magnetic stimulation (TMS) applied to the putative human homolog of MST, but not to putative human MT, selectively impaired performance in a discrimination task with radial and rotational motion patterns, providing further evidence for a division of the human MT+ complex that is similar to macaque MT/MST.

2.9 Conclusion

All of the information about the visual world that is available to an organism must be encoded in the neural responses that leave the two retinae. As David Marr (2010) pointed out, different representations of information will produce different aspects of reality. In that sense, one can think of the areas of visual cortex and their activity patterns as partial representations of the visual world, each emphasizing a facet. Information about edges, texture, color, and motion is all contained within the activity that reaches V1, but only after being processed by specialized brain areas with very specific connectivity patterns will this information be made explicit within the firing rates of individual neurons or small groups thereof. Correspondingly, much of modern cortical electrophysiology has been focused on identifying the distinguishing characteristics of the plethora of areas and their respective partial representations of the sensory environment. As we have reviewed here, these approaches have identified MST as a key area for integrating multimodal

information: information about motion, both linear and more complex, about heading, eye movements or memorized motion. All are explicitly encoded in its population activity. We argue that MST is more than just another processing step along the visual hierarchy that represents information about the visual world in yet another way. Once information reaches MST, it has undergone enough changes to be "ready for use". Especially with regard to self-motion perception, the evidence suggests that MSTd (likely in cooperation with surrounding areas, such as VIP) represents the information in a way that can be used by decision and motor areas to react. In other words, there is likely no need for further reshaping of its representations of visual features by downstream areas.

One aspect that needs to be kept in mind is that MST is composed of subregions, which we have alluded to multiple times throughout this review. Even within a subregion, individual neurons often show a wide array of behaviors: they respond better to the motion of individual objects or to the wide-field motion of the background; they do or do not respond to vestibular cues; they have congruent or incongruent preferred directions for visual and vestibular cues; they encode an optic flow fields focus of expansion in retinal or in real-world coordinates. As Bradley et al. (1996) pointed out, this could mean that only some MSTd neurons are actively involved in heading perception, while other MSTd neurons are contributing to other perceptual tasks. Alternatively, these differences in response patterns could be a sign that a lot of computation is happening within the area and that these different cell types represent different stages of these computations. It is probably because these complex response patterns in MST defy an easy description of its function that this area has received less prominence as a model system for sensory, cognitive, and motor planning processes than MT[2]. Comparatively simple approaches for characterizing a visual cortical area, such as tuning curves, work well in early and midlevel areas, such as MT or V4. Higher level areas, such as IT and MST, require more complex methods, such as adaptive sampling of complex stimuli (Yamane, Carlson, Bowman, Wang, & Connor, 2008) or hierarchical convolutional neural networks (Yamins & DiCarlo, 2016). But now, the time might be ripe to embrace the challenge and appreciate the role of MST at the intersection of sensation and cognition. The overarching goal should be to get a better understanding of MSTs role in everyday behavior. This requires the combination of different methodologies, such as psychophysics, functional imaging, electrophysiology and disruptive methods like TMS and optogenetics. They should be applied across different experimental paradigms

[2]A Google Scholar search for middle temporal lists 160,000 hits, compared to slightly more than 6,400 for medial superior temporal (as of Jan. 19, 2021)

that aim at simulating more natural behavior, such as free-viewing (Bremmer et al., 2017) and combined visual and vestibular stimulation in the face of self- and object motion (e.g., Sasaki et al., 2017). The combination of functional imaging, electrophysiological recordings, microstimulation, and artificial neural networks emerged as a fruitful approach for discovering new principles that govern the organization of IT cortex (Bao et al., 2020) and could be useful in advancing our understanding of MST. Recent advances in wireless recording techniques in freely moving animals (e.g., Berger, Agha, & Gail, 2020) in combination with improved behavioral tracking methods (A. Mathis et al., 2018) might open many new opportunities in the near future. It is precisely because of its unique position as a gateway between perception and cognition that we believe MST to be an ideal model system of a core feature of higher nervous systems: the transformation of sensation into multidimensional internal representations of a dynamic environment, enabling cognition and sophisticated action planning.

Chapter 3

How does the brain tell self-motion from object motion?

Benedict Wild

The Journal of Neuroscience, Vol. 38 (2018)

Chapter 3 is a short review of an article by Sasaki, Angelaki, and DeAngelis (2017) entitled "Dissociation of Self-Motion and Object Motion by Linear Population Decoding That Approximates Marginalization". In this paper, the authors show that information about self-motion and object motion can be decoded from a population of multisensory neurons in macaque cortical area MSTd with an algorithm that approximates the mathematical process of marginalization.

This review was published in the *Journal of Neuroscience*'s "Journal Club" section (Wild, 2018), which allows students or postdoctoral fellows to write scholarly reviews, without their respective supervisors, of papers that have recently been published in the journal.

Acknowledgements

I thank Prof. Stefan Treue and Dr. Pierre Morel for helpful comments and discussions on the manuscript; and Bill Galic for proofreading.

Imagine a football player who has to run across the field, evade opponent players, and eventually catch the ball that is passed to him by his teammate. As he does this, his retinal image is a hodgepodge of moving components. The position of background objects (e.g., yard lines, goal posts, the cheering crowd) changes relative to the runner as he runs across the field, creating a so-called "optic flow" pattern on his retina; the opposing team's players move toward him from different angles; and, most importantly, the football eventually appears in his field of vision. To navigate and act within this ever-changing environment, the player must be able to distinguish retinal motion resulting from his own movements from the motion caused by moving objects. But how can the relevant motion components be selected and confounding components be discarded, given that all the information the visual system receives is the jumble of movements across the retina?

The dorsal part of the medial superior temporal area in the macaque cortex (area MSTd) has long been known to play an important role in motion perception in general (Graziano et al., 1994) and selfmotion perception in particular (Britten, 2008). Individual neurons in this part of the brain are typically tuned for the direction of a moving stimulus or for a particular heading direction, as measured with stimuli that simulate optic flow (e.g., a field of random dots that move away from a single point, which determines the direction of the simulated self-motion). Furthermore, the area contains multimodal neurons that integrate visual information with vestibular information (Duffy, 1998, Gu et al., 2008) for a joint representation of selfmotion. Interestingly, some multimodal MSTd neurons have the same preferred direction for visual and vestibular input ("congruent cells"), whereas others prefer different or even opposing directions ("opposite cells") (Gu et al., 2008). During simultaneous visual and vestibular stimulation, the neuronal sensitivity for discriminating heading directions (calculated based on signal detection theory) (see Britten et al., 1992) is decreased in such opposite cells compared with either visual or vestibular stimulation alone. Congruent cells, on the other hand, show an increased sensitivity in this "bimodal condition" compared with either unimodal condition (Gu et al., 2008). This raises the question of what the purpose of opposite cells may be.

In a recent publication in The Journal of Neuroscience, Sasaki et al. (2017) suggested that opposite cells play an important role in parsing object motion and self-motion-components from the overall retinal image motion. To investigate this, they recorded the activity of individual neurons in area MSTd in 2 macaque monkeys who were placed in a virtual-reality setting. The monkeys were seated on a platform that could passively be moved in 3D space, thus providing vestibular stimulation. Simultaneously, visual optic flow stimuli were presented through a 3D field of stars that simulated translational

selfmotion in one of eight directions in the frontoparallel plane. Additionally, a cluster of nine spheres, defined by increased dot density (the "object"), moved in one of eight possible directions through the visual world on some of the trials (Sasaki et al., 2017, their Fig. 1).

Sasaki et al. (2017) found that the influence of object motion on heading tuning (and the influence of heading direction on object motion tuning) differed between cell types: for congruent cells (50 % of recorded cells), heading tuning was more consistent in the bimodal condition (vestibular and visual stimulation) than in the visual-only condition, but object motion direction tuning was more consistent in the visual-only condition. Conversely, for opposite cells (∼18% of recorded cells), heading tuning was stronger in the visualonly condition, whereas object-motiondirection tuning was more consistent in the bimodal condition. This makes intuitive sense: if a cells preferred heading direction is the same for visual and for vestibular stimuli (i.e., congruent cells), then adding vestibular information will help the cell to maintain its normal tuning curve in the face of visual input that is confounded by a moving object. For cells that have opposing preferences for visual and vestibular heading information (i.e., opposite cells), bimodal stimulation flattens the tuning curve, thereby decreasing the selectivity for heading direction. Furthermore, the preferred direction for moving objects typically is opposite to the preferred heading direction because selfmotion in one direction (e.g., to the left) means that the retinal image, including individual objects, moves in the opposite direction (e.g., to the right). A cell that is tuned for heading to the left should therefore also be tuned for an object moving to the right. Thus, in opposite cells, vestibular heading tuning is aligned with object motion direction tuning, as both are the opposite of visual heading tuning.

How do these different cell types influence the way that heading direction and object motion direction are represented by a population of MSTd neurons? To determine which specific stimulus most likely elicited a given neural population response, the authors first computed the joint probability of a specific heading and a specific object motion given the population response. Despite some simplifying assumptions, this decoder was able to accurately estimate both the heading and the object motion. However, this strategy would become computationally expensive if there were several objects moving through the scene (which is often the case in reallife scenes), as this would require the brain to calculate a multidimensional probability distribution.

This problem can be solved, however, by a mathematical procedure called marginalization, which determines the probability of one event (e.g., a specific heading direction) independent of a second event (e.g., object motion) that modulates the probability of the first event. While models have previously suggested that this process can be implemented in the brain (Beck et al., 2011),

Sasaki et al. (2017) considered an approximation of marginalization, which, they claimed, would be more intuitive in terms of neuronal implementation. To this end, they first tried to decode heading direction from the responses of bimodal neurons by calculating the likelihood of a given heading direction (or object motion direction) as the sum of each neuron's response, weighted by either its visual or its vestibular tuning curve (Jazayeri & Movshon, 2006). This approach was not successful, however, possibly because of the diversity of tuning properties across the neuronal population. Approximate linear marginalization (ALM) (Kim et al., 2016) differs from the traditional likelihoodcomputation in that it uses a regression model to find the optimal weights with which each neuron influences the overall likelihood of a specific heading direction. Applying this procedure to either a subset of the recorded MSTd neurons (all opposite cells and an equal number of randomly selected congruent cells) or to the whole sample resulted in a much better decoding performance. The decoding error was also smaller in the bimodal condition than in the visual-only condition, showing that the algorithm uses vestibular information to improve its decoding accuracy. The profile of the decoding weights across the population was similar to the neurons tuning curves, suggesting that the brain uses information it already has when determining how much each cell contributes to the populations representation of the stimulus. However, the ALM algorithm appeared to apply a gain factor to the neurons that was not inherent to the neurons' tuning properties.

In summary, Sasaki et al. (2017) showed that a new algorithm, which had previously been developed for a population of simulated MSTd neurons (Kim et al., 2016), can be applied to a population of real MSTd neurons to distinguish self-motion from object motion. These findings raise the question of whether such an algorithm can actually be implemented by the brain and, if so, how the brain could learn the weights that are assigned to each neuron. ALM learned the weights by being trained on 500 simulated trials for each stimulus condition, attempting to minimize the difference between the true probability distribution and the algorithms estimation of this distribution. During development, the brain does not have direct access to the true probability distribution to quickly learn the correct decoding weights. Instead, it acts based on its own estimation and then has to infer correct and incorrect judgments based on feedback to the actions it took. This should in principle allow decision-making areas in the brain to learn how to optimally read out the population response of MSTd.

Another point in question is why the authors put so much emphasis on marginalization being implemented through a linear decoder, at the cost of being only an approximation. They state that, for the purpose of modeling, they assume the brain to be "limited to processing neural responses linearly"

(Kim et al., 2016) and that nonlinear transformations "may be difficult for the brain to implement" (Sasaki et al., 2017). However, there is evidence that marginalization can be implemented in neural circuits through divisive normalization (Beck et al., 2011), a widespread nonlinear computation where neuronal responses are inhibited, and thus effectively normalized, by the summed activity of a pool of neurons (for review, see Carandini & Heeger, 2012). Furthermore, area MST has been suggested to integrate its visual input from the middle temporal area in a nonlinear manner (Mineault et al., 2012), raising additional doubts as to why decoding of the output of MST would have to be a linear approximation. Thus, although ALM can decode heading and object motion from a population response in area MSTd, this does not guarantee that the brain actually implements this specific computation.

What additional strategies could the brain use to distinguish self-motion from object motion? In many cases, self-motion is caused by actions of the individual, such as walking or running, and these actions provide extraretinal information, such as the stimulation of proprioceptive sensors in moving body parts, or efference copies of motor commands to other parts of the brain. These can be used to compensate for the effects of self-motion on the retinal image flow (e.g., Crowell, Banks, Shenoy, & Andersen, 1998), so that any uncompensated retinal motion is likely due to moving objects (Wallach, 1987). But even when being moved passively (e.g., while riding a train and looking out of the window), eye movements can provide additional information about object motion during self-motion (P. A. Warren & Rushton, 2007). This is particularly interesting, as MSTd neurons also carry signals about eye movements (Newsome et al., 1988).

In conclusion, Sasaki et al. (2017) provided evidence that area MSTd can help our football player achieve his goal. Information that is encoded by MSTd neurons can be used to compute the players own movement running across the field, even as his perceived optic flow patterns are disrupted by the movements of other players and the football. Similarly, MSTd can represent the balls trajectory, even though its motion on the retina is distorted by the players own movements. This can be achieved by a neuronal population of congruent and opposite multisensory cells, whose responses can be decoded in a way that approximates the mathematical process of marginalization to accurately estimate a single variable of interest. Whether this is the computation actually performed by the brain needs to be investigated in more detail. Showing that decision-related brain areas higher up in the cortical processing hierarchy represent heading or object motion information in a manner that is consistent with the read-out predicted by ALM would provide evidence in favor of this hypothesis.

Chapter 4

Electrophysiological dataset from macaque visual cortical area MST in response to a novel motion stimulus

Benedict Wild, Amr Maamoun, Yifan Mayr, Ralf Brockhausen, Stefan Treue

under review at Scientific Data

Chapter 4 describes a published comprehensive dataset from a neurophysiology study in macaque monkey visual cortex that includes a complete record of extracellular action potential recordings from the extrastriate medial superior temporal (MST) area, behavioral data, and detailed stimulus records. Analyses of the data described here are presented in chapter 5.

Author contributions

A.M. and S.T. conceived the experiment. A.M. and B.W. collected the data. Y.M., R.B., B.W., and S.T. developed and implemented the dataset structure. B.W. and Y.M. wrote the example code. B.W. wrote the manuscript. S.T. supervised the project. All authors reviewed the manuscript.

Acknowledgements

This work was supported by grants to S.T. from the Deutsche Forschungsgemeinschaft (DFG) through the Research Unit 1847 “Physiology of Distributed

Computing Underlying Higher Brain Functions in Non-Human Primates" (Project A1) and the Collaborative Research Center 889 "Cellular Mechanisms of Sensory Processing" (Project C04). We thank Leonore Burchardt, Sina Plümer, and Dirk Prüsse for assistance in animal handling, training, and surgeries, and the German Primate Center's veterinary and animal husbandry staff for their expert animal care.

Abstract

Establishing the cortical neural representation of visual stimuli is a central challenge of systems neuroscience. Publicly available data would allow a broad range of scientific analyses and hypothesis testing, but are rare and largely focused on the early visual system. To address the shortage of open data from *higher* visual areas, we provide a comprehensive dataset from a neurophysiology study in macaque monkey visual cortex that includes a complete record of extracellular action potential recordings from the extrastriate medial superior temporal (MST) area, behavioral data, and detailed stimulus records. It includes spiking activity of 172 single neurons recorded in 139 sessions from 4 hemispheres of 3 rhesus macaque monkeys. The data was collected across 3 experiments, designed to characterize the response properties of MST neurons to complex motion stimuli. This data can be used to elucidate visual information processing at the level of single neurons in a high-level area of primate visual cortex. Providing open access to this dataset also promotes the 3R-principle of responsible animal research.

4.1 Background & summary

Determining and quantifying the relation between physical stimuli and the neuronal responses they evoke is the most widely taken approach in sensory neuroscience, resulting in a variety of modeling and analysis approaches (Paninski, Pillow, & Lewi, 2007; Schwartz et al., 2006; Wu, David, & Gallant, 2006) for determining the neural representation of stimulus parameters. Most studies focus on early, low-level visual areas, such as the retina (L. D. Liu & Pack, 2017; Maheswaranathan, Kastner, Baccus, & Ganguli, 2018; Pillow et al., 2008), lateral geniculate nucleus (LGN) (Cai, Deangelis, & Freeman, 1997; Dan, Alonso, Usrey, & Reid, 1998; Solomon, Tailby, Cheong, & Camp, 2010), or primary visual cortex (V1) (Jones & Palmer, 1987; Park & Pillow, 2011; Rust et al., 2005; Touryan et al., 2002) of the mammalian brain. In primates, however, the visual system is hierarchically structured (Felleman & Van Essen, 1991) with neurons in "higher" areas – downstream of V1 – progressively showing more complex stimulus preferences (Treue, 2003). For example, stimulus representations tend to become more "tolerant" or "invariant" for position, scale, and context from mid- to high-level areas in the ventral (Rust & DiCarlo, 2010) and dorsal (Wild & Treue, 2021b) pathways of the primate cortex. It has been proposed that conscious visual perception begins at the top of the processing hierarchy and that the quick, categorical recognition of objects ("forest before trees") that guides most of our behavior relies predominantly on activity in the higher areas (Hochstein & Ahissar, 2002). Therefore, a better understanding of the functional relationship between stimulus features and neuronal responses in mid- and high-level areas is essential for our understanding of sensation and brain function more broadly. Nevertheless, much fewer studies have focused on stimulus representations in mid-level or higher visual areas as compared to the rich literature on neural encoding of stimulus features in the early visual system (but see Rust et al., 2006; Yamane et al., 2008; Rust & DiCarlo, 2010; Mineault et al., 2012 for examples). Even fewer such studies provide access to their data. For example, the Collaborative Research in Computational Neuroscience (CRCNS) website (`crcns.org`) offers 13 data sets with recordings from V1, but only 3 data sets of MT recordings (Cui et al., 2013; Nishimoto & Gallant, 2018; Niknam et al., 2018), one of V4 recordings (M. Smith, 2020), and one of MST/VIP recordings (Gu et al., 2018)

To address this shortage we provide a freely accessible and well curated comprehensive dataset from a neurophysiology study in macaque monkey visual cortex that includes a complete record of extracellular action potential recordings from the extrastriate medial superior temporal (MST) area, behavioral data, and detailed stimulus records. MST is a key area in the

dorsal visual pathway that receives its major input from motion-sensitive area MT (Born & Bradley, 2005) and is involved in the processing of complex motion as well as self-motion perception (Graziano et al., 1994; Mineault et al., 2012; Saito et al., 1986; Tanaka et al., 1986; Wild & Treue, 2021b). Notably, MST neurons have a number of complex features, such as position invariance (Graziano et al., 1994), and modeling work suggests that they perform a nonlinear integration of the output of MT neurons (Mineault et al., 2012). Models that include a nonlinearity are typically much more difficult to fit to data and require iterative procedures (as compared to, for example, simple linear-nonlinear (NL) cascade models for which the spike-triggered average corresponds to the maximum likelihood estimator of the neuron's spatiotemporal receptive field (Pillow, 2007). We hope this dataset, which contains neuronal responses to different well-parameterized motion stimuli, can serve as a starting point to develop such models.

We provide the spiking activity of 172 MST neurons that were recorded across three different experiments. The datasets from the first two experiments ("Spatial Mapping" and "Tuning") result from an elaborate and more systematic version of previous approaches (similar to, e.g., (Graziano et al., 1994; Saito et al., 1986; Tanaka et al., 1986)) and can be used to determine the neurons' spatial receptive fields as well as their tuning for direction and speed. For the "Spatial Mapping" experiment, small random dot patterns (RDPs) were sequentially presented in different locations across the screen, thus probing a neuron's response to random motion across the visual field. The "Tuning" experiment presented RDPs that were moving at different speeds in different directions in linear motion and spiral motion space in multiple locations across each neuron's spatial receptive field. As these two experiments take the standard approach for acquiring the stimulus-response function they can be considered the "ground truth" about visual response properties of MST neurons. The third experiment ("Reverse Correlation") recorded spiking responses to a newly developed random dot motion stimulus which consists of a grid of positions, each with a direction and speed seed that determine the motion of dots in the vicinity of each grid location. Each seed has a Gaussian weighting field that determines its influence on the motion of surrounding dots. This creates an overall smooth, wave-like motion pattern. Each seed was assigned a new direction and speed every 100 ms, thus creating a temporal structure of the stimulus that makes it well-suited for spike-triggered analysis approaches (Schwartz et al., 2006). We believe that providing this dataset to the scientific community will be of high value to theoretical, computational, and systems neuroscientists who are interested in visual processing beyond retina, LGN, and V1.

4.2 Methods

Animal welfare statement

Research with non-human primates represents a small but indispensable component of neuroscience research (Buffalo et al., 2019; Treue & Lemon, 2021). The scientists in this study are aware and are committed to the responsibility they have in ensuring the best possible science with the least possible harm to the animals (Roelfsema & Treue, 2014). Providing the data collected in this study in a well-curated format and with open access to the scientific community contributes to this commitment. Such datasets ensure maximal transparency about study results and promote the accessibility of data, ensuring its best use. This is in line with the 3R-principle and international efforts to improve the reporting an accessibility of biomedical research data (Percie du Sert et al., 2020).

All animal procedures of this study have been approved by the responsible regional government office (Niedersaechsisches Landesamt fuer Verbraucherschutz und Lebensmittelsicherheit, LAVES) under the permit numbers 3392 42502-04-13/1100 and 33.19-42502-04-18/2823. The animals were group-housed with other macaque monkeys in facilities of the German Primate Center in Goettingen, Germany in accordance with all applicable German and European regulations. The facility provides the animals with an enriched environment (including a multitude of toys and wooden structures (Berger et al., 2018; Calapai et al., 2016)), natural as well as artificial light, exceeding the size requirements of the European regulations, and access to outdoor space. We have established a comprehensive set of measures to ensure that the severity of our experimental procedures falls into the category of mild to moderate, according to the severity categorization of Annex VIII of the European Unions directive 2010/63/EU on the protection of animals used for scientific purposes (Pfefferle et al., 2018). Surgeries were performed aseptically under gas anesthesia using standard techniques, including appropriate perisurgical analgesia and monitoring to minimize potential suffering. The German Primate Center has several staff veterinarians who regularly monitor and examine the animals and consult on procedures. During the study, the animals had unrestricted access to food and fluid, except on the days where data were collected or the animal was trained on the behavioral paradigm. On these days, the animals were allowed unlimited access to fluid through their performance in the behavioral paradigm. Here the animals received fluid rewards for every correctly performed trial. Throughout the study, the animals' psychological and veterinary welfare was monitored by the veterinarians, the animal facility staff and the labs scientists, all specialized in working

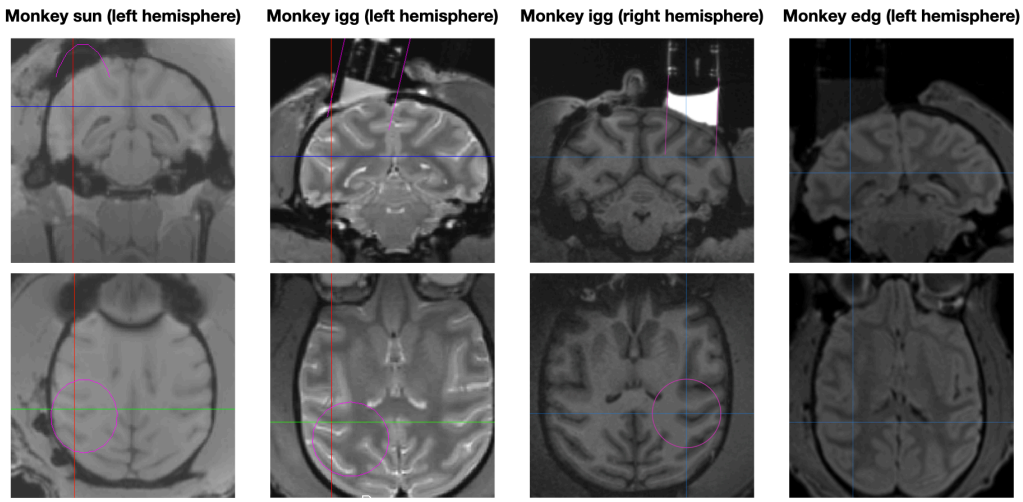

Figure 4.1: Implant locations of the recording chambers. The figure shows coronal (top row) and axial (bottom row) slices of MRI scans, displaying the location of the recording chamber in each of the 3 monkeys. Monkey igg was already implanted with a recording chamber over the left hemisphere when he joined the project (second column), but was re-implanted with another chamber over the right hemisphere (third column) during his participation in the project.

with non-human primates. The three animals were healthy at the conclusion of our study and were subsequently used in other studies.

Animals

Three male rhesus macaque monkeys (*Macaca mulatta*) contributed to this dataset. All animals had previously participated in other projects and were implanted with a titanium head holder to minimize head movements during the experiment, as well as with a recording chamber implanted over the parietal lobe based on a magnetic resonance imaging (MRI) scan. The recordings chambers were cylindrical with a diameter of 24 mm. Surgeries were conducted under general anesthesia and post-surgical care using standard techniques.

Monkey sun

Monkey sun was 14 - 16 years old and weighed between 9.9 and 12.8 kg during the period of data collection. The monkey had previously participated in other projects and was already implanted with a head-post and a chamber

over the left hemisphere (stereotactic coordinates: mediolateral (ML): 15 mm left; anteroposterior (AP): 11 mm posterior, Fig. 4.1). The chamber was angled posterior with an inclination of 30° from the vertical. 82 of the 138 (59%) recordings included in this dataset come from Monkey sun.

Monkey igg

Monkey igg was 10 - 13 years old and weighed between 10.4 and 13.4 kg during the period of data collection. The monkey had previously participated in other projects and was already implanted with a head-post and a chamber over the left hemisphere (coordinates: ML: 8 mm left; AP: 5 mm posterior, Fig. 4.1). The chamber was angled medial with an inclination of 18° from the vertical. During his participation in the project, the recording chamber over the left hemisphere was removed and a new vertically oriented chamber was implanted over the right hemisphere (coordinates: ML: 17 mm right ; AP: 1 mm anterior, Fig. 4.1). 54 of the 138 (39%) recordings included in this dataset come from Monkey igg (26 from the left and 28 from the right hemisphere).

Monkey edg

Monkey edg was 14 years old and weighed between 8.0 and 9.2 kg during the period of data collection. The monkey had previously participated in other projects and was already implanted with a head-post and a vertically oriented chamber over the left hemisphere (coordinates: ML: 13 mm left; AP: 0 mm, Fig. 4.1). 2 of the 138 dataset (2%) recordings included in this dataset come from Monkey edg.

General experimental setup

Data were collected from two different experimental setups (setup A: Monkeys sun and edg as well as some data from Monkey igg; setup B: most data from Monkey igg). In both setups eye position was recorded binocularly with an Eyelink 1000 system (SR-Research, Ottawa, ON, Canada) at a sample rate of 500 Hz. The experiments were controlled by the open-source software MWorks (`mworks-project.org`) running on two Apple Macintosh computers (Apple Inc., Cupertino, CA, USA), a client and a server.

Setup A

During recordings a monkey was seated in a custom-built primate chair and viewed a 27" LCD monitor (XL2720T, BenQ, Taipei, Taiwan) from a

Stimulus Type	**Setup A**	**Setup B**
Background	17.3 cd/m^2	1.0 cd/m^2
White dots	82.7 cd/m^2	31.3 cd/m^2
Fixation point (bright)	13.4 cd/m^2	1.9 cd/m^2
Fixation point (dim)	8.9 cd/m^2	0.6 cd/m^2

Table 4.1: Luminance values of all stimuli in the two recording setups.

fixed distance of 57 cm in a dark room. The monitor had a resolution of 1920×1080 pixels, a refresh rate of 120 Hz, and covered $60° \times 30°$ of the visual field. The luminance values of the background and the different stimuli are listed in Table 4.1.

Setup B

The conditions in this setup were the same as in setup A, except that a 171.5×107.2 cm back projection screen (dnp Black Bead, Karlslunde, Denmark) was viewed from a fixed distance of 102 cm. Stimuli were displayed via a projector (projection design, Fredrikstad, Norway) with a a resolution of 1920×1200 pixels and a refresh rate of 60 Hz.

Task

Throughout all three experiments, monkeys performed a simple luminance change detection task. At the beginning of each trial a red fixation square (size: $0.2° \times 0.21°$) appeared. To make sure the different stimuli described below did not occlude the fixation square, a circular "mask" with a diameter of 1.5° of the same color as the background surrounded the fixation square. Monkeys could initiate the stimulus presentation by keeping their gaze within a square window of side length 3° around the fixation point ("fixation window") and pressing a button attached to their primate chair. The luminance of the fixation point changed within a time window that was randomly selected from a uniform distribution ranging from 3 to 4.6 s after stimulus onset. The monkey had to respond to this luminance change within 600 ms ("reaction time window") in order to receive a fluid reward (juice, tea, or water, depending on each monkey's preferences). Monkeys sun and edg initiated a trial by briefly pressing the button and responded by pressing it again; Monkey igg kept the button pressed throughout the duration of the trial and responded by releasing it (these were the response patterns that the respective monkeys had been trained on in previous experiments). The exact luminance values before and after the change are listed in Table 4.1 for each setup. A trial

could end in one of three ways: with a juice reward and a distinct sound if the monkey kept its gaze within the fixation window throughout the trial and responded to the luminance change within the reaction time window; with no reward and a different sound signaling an error if the monkey responded too early or too late (i.e., outside the reaction time window); or with no reward and a third sound if the monkey's gaze moved outside the fixation window during the trial. The position of the fixation point was optimized in each recording session to ensure that as much of the receptive field as possible was covered by the display (within the constraints that fixation points too far out in the periphery made it difficult for the monkey to keep its gaze within the fixation window for a prolonged period of time).

Experiments

On a given day the animal had to complete the experiments in the order "Spatial Mapping", "Tuning", "Reverse correlation". The sum of all recordings on a given day are called one "recording session". Given that the animal determined the number of trials performed each day, not all recording sessions contained complete recordings for all three experiments. Table 4.2 shows the number of recording sessions as well as the number of recorded neurons for each of the three experiments.

	Spatial Mapping	Tuning	Reverse Correlation
Number of sessions:	139	139	119
Number of recorded neurons:	172	172	150

Table 4.2: Numbers of recording sessions and recorded units for each of the three experiments.

Experiment 1 (Spatial Mapping)

In Experiment 1, a small random dot pattern (RDP) was presented sequentially in different locations. The RDP consisted of white dots (cf. Table 4.1 for luminance values) with a radius of 0.2° that were randomly placed in a circular aperture with a radius of 1.5°. The dot density was 4 dots/deg^2. Dots moved independently along linear trajectories in random directions and dots that moved out of the aperture were replotted at a random location within the aperture. Each presentation of the RDP for 50 ms constitutes one "sample". For every sample, the location of the RDP was updated by assigning a random x-position in the range covering the entire width of the display

(-28° to 28° in setup A, -30° to 30° in setup B) in steps of 2° and a random y-position in the range covering the entire height of the display (-16° to 16° in setup A, -20° to 20° in setup B) in steps of 2°, where the location [0°, 0°] is the center of the screen. The speed of the dots was randomly chosen from a range of 4°/s 24°/s for every sample. On trials that were not interrupted by fixation breaks or incorrect responses (see Task), more than 80 samples could be presented. In 76 cells, 10% of the sample presentations were blank (i.e., no stimulus other than the fixation point was shown on the screen for 50 ms). The purpose of these "blank presentations" was to calculate the baseline firing-rate of the cell in this task and to explore the cells' responses to the random and quick location changes of the RDP on a longer timescale.

Experiment 2 (Tuning)

In Experiment 2, an RDP was presented sequentially in up to 5 different locations overlapping with the spatial receptive field as it had been determined by an online analysis of Experiment 1. The RDP consisted of white dots (cf. Table 1 for luminance values) with a radius of 0.2 that were randomly placed in a circular aperture with a radius ranging from 5° to 10°, depending on the size of the spatial receptive field. The dot density was 2 dots/deg^2. The location, direction, and speed of the RDP were updated every 100 ms, the time period that constitutes one "sample" in Experiment 2. For most of the cells, the 5 locations were arranged in a cloverleaf, similar to a previous study (Graziano et al., 1994). In a minority of cells whose receptive field was on the edge or corner of the screen, the arrangement was different to ensure maximal coverage of the receptive field by the 5 RDP positions. For the position overlapping with the center of the receptive field, direction was chosen randomly from a set of 8 different translational and 8 spiral directions (see Table 4.3). For the other 4 locations, direction was randomly chosen from the 8 spiral directions. "Spiral space" is a one-dimensional space that includes radial, rotational, and spiral motion patterns. A direction in spiral space can be specified in degrees or radians, similar to translational directions. For an RDP that is moving in direction α in spiral space, the displacement of each dot is determined by updating its polar coordinates r (radius) and θ (angle) (with respect to the RDP's center) as

$$r = r + s * \cos(\alpha)$$

$$\theta = \theta + s * \sin(\alpha)$$

where s is a constant that depends on speed. Thus, for $\alpha = 0°$, the distance of each dot from the RDP's center (i.e., it's radius, r) increases maximally

(because $\cos(0) = 1$) and the angle does not change at all (because $\sin(0) = 0$), which results in expansion (all dots moving away from the center). For $\alpha = 90°$, each dot's radius remains the same (because $\cos(90) = 0$) and the angle increases maximally (because $\sin(90) = 1$), which results in clockwise rotation. Table 4.3 shows an overview of the 8 directions in translational and spiral space.

Direction in degrees	**Linear Motion**	**Spiral Motion**
0°	Up	Expansion
45°	Diagonal up and right	Clockwise outward spiral
90°	Right	Clockwise rotation
135°	Diagonal down and right	Clockwise inward spiral
180°	Down	Contraction
225°	Diagonal down and left	Counterclockwise inward spiral
270°	Left	Counterclockwise rotation
315°	Diagonal up and left	Counterclockwise outward spiral

Table 4.3: Definitions of translational and spiral motion directions. Note that the definition for translational motion deviates from the geometric convention where 0° is rightward and angle increases in counterclockwise direction; instead 0° is defined was upward motion and angle increases in the clockwise direction.

The speed of each RDP was randomly selected for every sample from one of six discrete values evenly spaced from 4°/s 24°/s. During spiral motion, RDPs had a speed gradient with the specified speed determining the speed of dots at a distance of 1° from the RDP's center. As in Experiment 1, 10% of samples were "blank samples" (i.e., no RDP was shown) in 76 cells.

Experiment 3 (Reverse Correlation)

In Experiment 3, a newly developed large, rectangular RDP was presented, which we call RC stimulus ("reverse correlation stimulus"). What the monkey sees is a smooth, wave-like pattern of moving dots (a video of an example trial is provided in the repository, see below for detailed information). Technically, the stimulus consists of 10×15 square segments, each with a side length of 3°, resulting in an overall size of $30° \times 45°$. Each segment is assigned a random translational direction and a random speed from 0°/s 20°/s every 100 ms, the time period that constitutes one "sample" in Experiment 3. The

center of each segment provides a Gaussian weighting field which determines its influence on the dots. Each dot's direction and speed are calculated as

$$f(x,y) = \frac{1}{n} \frac{\sum_i^n p_i \exp(-(\frac{(x-x_i)^2}{2\sigma_1^2} + \frac{(y-y_i)^2}{2\sigma_2^2}))}{\sum_i^n \exp(-(\frac{(x-x_i)^2}{2\sigma_1^2} + \frac{(y-y_i)^2}{2\sigma_2^2}))}$$

with the following variables:

n	number of neighboring segments ranging from 3 to 9
p_i	parameter (direction or speed) in question of segment
x, y	coordinates of the dot in degrees relative to the fixation point
x_i,y_i	coordinates of center of segment i relative to the fixation point
σ_1	standard deviation of the Gaussian filter along its first axis
σ_2	standard deviation of the Gaussian filter along its second axis

The standard deviations of the Gaussian filter σ_1 and σ_2 were both set to a value of 1.2° in all experiments. Each segment contained 10 dots with a radius of 0.2° that were randomly placed inside the segment (cf. Table 1 for luminance values), resulting in a dot density of 1.1 dots/deg^2 and a total number of 1500 dots for the entire stimulus. For 47 recording sessions conducted with Monkey igg, parts of the RC stimulus were masked with rectangles of the same color and luminance as the background, so as to effectively reduce the size of the stimulus to 6×9 square segments (overall stimulus size of $18° \times 27°$).

Neural recording setup

For each recording session, between one and three microelectrodes (Thomas Recording, Giessen, Germany) were advanced into the bank of the superior temporal sulcus, targeting area MST using microdrives. The microdrive was mounted at the beginning of each session onto the recording chamber and a x-position (on the medio-lateral axis) and a y-position (on the anterior-posterior axis) of the single or central electrode were determined. For most of the recording sessions, electrode depth position was controlled with a 3 electrode or a 5 electrode "Mini Matrix" system (Thomas Recording, Giessen, Germany), where electrodes are advanced using a rubber tube mechanical system. For a small subset of recordings from the left hemisphere of Monkey igg (a total of 22 recordings), we used the Model 650 single electrode Micropositioner (David Kopf Instruments, Tujunga, CA, USA), where the electrode is advanced using a hydraulic system. Signals from the electrodes were amplified and recorded with a sampling rate of 40 kHz and 16-bit precision using an Omniplex acquisition system (Plexon, Dallas, TX, USA).

Data preprocessing

Action potentials ("spikes") were identified using OfflineSorter V4 (Plexon, Dallas, TX, USA). The raw data was filtered with a 6-pole Bessel high-pass at a cut-off frequency of 250 Hz and spike waveforms were detected based on a manually determined threshold. These waveforms were then manually split into clusters based on different features as implemented in the software, including the first three principal components of the waveforms, the maximum and minimum voltage amplitude across the entire waveform length ("peak" and "valley"), or the waveform energy. For each recording, features were chosen according to the best separation between clusters. Note that for the large majority of recordings, there was only unit recorded so that this procedure predominantly served the purpose to separate the signal from background noise. Only in 29 sessions 2 units and in 2 sessions 3 units were recorded simultaneously, which required actual sorting of waveforms as belonging to different units. Behavioral and stimulus parameters of each recording session were originally stored in the MWK format as a stream of MWorks events (`mworks-project.org`). The lists contain the time, the name of the event, and data associated with the event. To increase data accessibility, the relevant parameters of each experiment, together with online and offline sorted spikes, were converted into the HDF5 format (see Usage Notes).

4.3 Data records

All data and supplementary material are publicly available via the German Neuroinformatics Node (G-Node, `http://www.g-node.org`). We provide a complete set in a permanently archived format ($\sim$ 27 GB), and in addition maintain a repository[1] for possible future updates of the data and the supplementary material (Wild, Maamoun, Mayr, Brockhausen, & Treue, 2021). The datasets are provided in the Hierarchical Data Format 5 (HDF5). We chose the HDF5 format because it is a portable and self-describing file format where data and metadata can be passed along in one file, in accordance with the FAIR guiding principles for scientific data management (Wilkinson et al., 2016). The file structure of HDF5 files includes only two major types of object - datasets and groups. Datasets are homogeneous n-dimensional arrays, and groups are container structures which hold datasets and/or groups. Metadata can be added to datasets and groups as attributes. The HDF5 data model, file format, API, library, and tools are open and distributed with-

[1] this repository will be made public once we have addressed potential reviewer comments

out charge (`https://www.hdfgroup.org/solutions/hdf5`) and the content of HDF5 files can be directly viewed in the freely available HDF Viewer (`https://www.hdfgroup.org/downloads/hdfview/`). Furthermore, the format can easily be converted and is therefore accessible using multiple widely used programming languages, such as MATLAB, Python, R, and Julia. Because of these advantages this format has seen increased popularity in recent years for the storage of neuroscience data (Brochier et al., 2018; Diggelmann, Fiscella, Hierlemann, & Franke, 2018; Herz, Meier, Nawrot, Schiegel, & Zito, 2008; Zehl et al., 2016; Stoewer, Kellner, Benda, Wachtler, & Grewe, 2014).

Data from each of the three experiments are organized in three folders: *MSTm* contains data from Experiment 1 ("Spatial Mapping"), *MSTt* contains data from Experiment 2 ("Tuning"), and *MSTn* contains data from Experiment 3 ("Reverse Correlation"). For each recording session there are two files: a 'task.h5' file contains the stimulus parameters, online and offline sorted spikes, and trial descriptions; an 'eye.h5' file stores gaze position and pupil size. For some recording sessions, no data is available for Experiment 3 because of technical issues during the recording, because the recorded neuron was lost partway through the recording session, or because the monkey would not do enough trials on that day. Metadata about each recording session is provided as a separate tab separated value (tsv) file. Information is provided in a table, where each row describes one recording session and the columns are explained in Table 4.4. An additional table (data_description) provides definitions of all the variables. Lastly, the repository contains three videos in the .mp4 format of example trials from each of the three experiments (Exp1_spatial_mapping.mp4, Exp2_tuning.mp4, and Exp3_reverse_correlation.mp4).

Column heading	**Description**
format	Format version of the meta data table ("1.0" in call cases in this project)
recording_session	A running index of recording sessions
experimenter	The experimenter who recorded the data ("amm" for A.M. or "bew" for B.W.)
date	The date of the recording
monkey	The monkey ("sun", "igg", or "edg")
hemisphere	The hemisphere from which the data was recorded ("left" or "right")
chamber	A numerical identifier of the recording chamber on a hemisphere ("1" in all cases in this project)

session_number	A running counter of recording sessions for a given monkey (recorded as strings to include leading zeros which are needed to specify filenames)
daily_count	A running counter of separate recording attempts on a given day
hardware	The recording setup in which the file was recorded ("A" or "B") and the micropositioner that was used for the recording session (either "MM1" for the 5 electrode mini matrix, "MM2" for the 3 electrode mini matrix or "Kopf" for the hydraulic Model 650 Micropositioner, see Neural recording setup)
offline_units_n	n (here: $n = 3$) columns with an identifier for up to n offline-sorted single cells. The digits before the decimal point specify the electrode while digits after the decimal point specify different offline sorted units recorded on the same electrode. For example, 34.1 and 34.2 are two units recorded on the same electrode that were sorted offline, whereas 34.1 and 35.1 are two units that were recorded on separate electrodes
ML_n	n (here: $n = 3$) columns with the stereotaxic coordinates on the mediolateral axis of the electrode for up to n offline-sorted single cells that were identified in the columns "offline_units_n" (in mm)
AP_n	n (here: $n = 3$) columns with the stereotaxic coordinates on the anteroposterior axis of the electrode for up to n offline-sorted single cells that were identified in the columns "offline_units_n" (in mm)
depth_n	n columns (here: $n = 3$) with depth positions of the electrode for up to n offline-sorted single cells that were identified in the columns "offline_units_n", where 0 is the surface of the dura mater (in μm)
Exp_MSTm_tt, Exp_MSTt_tt, Exp_MSTn_tt	One column for each experiment (MSTm, MSTt, MSTn) that specifies the total number trials in this file (tt = total trials)
Exp_MSTm_st_n, Exp_MSTt_st_n, Exp_MSTn_st_n	n columns (here: $n = 3$), for each combination of experiment (MSTm, MSTt, MSTn) and up to n units that specify the range of trials on which spikes were recorded (st = spike trials)

notes	Additional information about a recording session that is not recorded in any of the previous columns.

Table 4.4: Description of the meta data as provided in a separate tsv file

4.4 Technical validation

We performed a number of plausibility checks to ensure the quality of the data and to rule out some potential problems that could bias the analysis of the data. First, we verified that the following is true:

- There are no negative spike times.
- Trial end times are always later than trial start times, i.e., all trials have a positive duration.
- The number of timestamps for every event is identical to the number of event values (see Usage Notes for a detailed description of event values and timestamps).

Trial duration

Individual trials can be arbitrarily short, because a trial was aborted immediately if the monkey's gaze exited the fixation window or the monkey made a premature response (see Methods). Trials can also be arbitrarily long, when the monkey pressed a button in response to the luminance change of the fixation point (see Methods) and then kept it pressed, rather than releasing it quickly. We therefore checked the time from trial start until the assignment of a trial outcome (hit, failure, or fixation break) in "hit"-trials, as these were not interrupted by fixation breaks or premature responses. The durations of these trials ranged from 3112 ms to 4604 ms (mean ± SD: 3832 ± 313 ms), which is in accordance with the values that had been specified (3.0 to 4.6 s).

Firing rate

In each of the three experiments, trials consist of rapidly changing presentations of randomly selected stimuli. On any given trial this includes stimuli that drive the cell's spiking activity strongly as well as stimuli that drive it hardly at all, so that the firing rate within a trial is expected to remain stable on average, without any systematic deviation over time. We do expect

the firing rate per trial to be higher in the "Tuning" experiment than in the "Spatial Mapping" experiment, because all stimuli were presented inside the receptive field and the stimuli were typically larger and contained coherent motion (compared to the small, incoherent motion RDPs used for "Spatial Mapping", see Methods). And indeed, a paired t-test showed a highly significant difference in the average firing rate of "Spatial Mapping" trials (16.11 spikes/s) and "Tuning" trials (19.91 spikes/s) ($t(138) = 6.21, p < .001$). Across all recordings and all experiments, 1972 trials that last at least 500 ms have 0 spikes. Across all units, all experiments, and all trials that last at least 500ms, the average firing rate (calculated per trial) is 16.60 spikes/s (SD: 13.30 spikes/s) and the highest firing rate calculated for one individual trial is 162 spikes/s. This value seems quite high and came from a trial that was quite short (833 ms) but contained a very large number of spikes (135). As we note in the next paragraph, some of these outliers might be caused by incorrectly classified waveforms during spike sorting, but across the entire data set, they are very rare.

Interspike intervals

The longest interspike interval across all files is less than 4 s, which we consider realistic and plausible. Across all recordings and experiments, there were 6996 interspike intervals (0.07%) that were longer than 2 s. 12216 interspike intervals (0.11%) were shorter than 1 ms. Because of small shifts in the position of the electrode over time, the quality of the voltage signal can change drastically. If the waveform changes too much, spike sorting might miss some spikes, misclassify them as belonging to a different unit, or incorrectly classify waveforms as belonging to the unit event though they do not. While it is nearly impossible to rule out misclassifications for sure, we believe that the information we provide here about firing rates and interspike intervals strongly suggests that this issue is negligible in our dataset. However, it should be kept in mind that each individual spike has a certain, albeit very low, probability of being misclassified which might be relevant for certain types of analyses.

In summary, we believe that these tests provide ample evidence for the quality of the data.

4.5 Usage notes

Our dataset is organized by experiment. Within each experiment folder there are two files for every recording session. Files ending in '-task.h5' contain

information about the stimuli, behavior, and electrophysiological recordings; files ending in '-eye.h5' contain eye tracking data. Each of these file pairs contains the complete data from all the trials the monkey completed. Here, we provide a general overview of this data structure and how variables of interest can be accessed. More specific examples for how this can be achieved in MATLAB are given by the example code (see Code availability). In both files the data is structured as a series of "events". There are five types of events that can be recognized by their names (the first four types only occur in -task.h5 files and the last type only occurs in -eye.h5 files):

- events that describe stimulus features (e.g., the direction or speed of a random dot pattern) start with STIM_
  There are 10 different stimuli with distinct prefixes:
  - STIM_background describes features of the background (all three experiments)
  - STIM_fixationPoint describes features of the fixation square (all three experiments)
  - STIM_MappingProbe describes features of the RDP in the Spatial Mapping experiment (Experiment 1 only)
  - STIM_mask describes features of the circular mask around the fixation square (all three experiments)
  - STIM_nDimRDP describes features of the RC stimulus (Experiment 3 only)
  - STIM_RecmaskBottom, STIM_RecmaskLeft, STIM_RecmaskRight, and STIM_RecmaskTop desribe features of the four rectangular masks that covered part of the RC stimulus in some of the recordings as described in the section Experiment 3 (Reverse Correlation) (Experiment 3 only)
  - STIM_TuningProbe describes features of the RDP in the Tuning experiment (Experiment 2 only)
- events that describe input/output variables (e.g., button press or reward delivery) start with IO_
- events that describe trial parameters (such as trial start or trial type) start with TRIAL_
- events that relate to action potential recordings start with SPIKE_
- events that describe eye tracking data start with EYE_

More detailed definitions of all events can be found the file data_description in the repository. Each event has a value and a timestamp, specifying the time in microseconds since the MWorks server was started on the day of the recording session. Each task-file (i.e., files ending in _task.h5) includes two structures – event_value, and event_time – of equal length that contain a field of values or timestamps for every event. As a concrete example, in the recording session "amm-MSTm-sun-120-01+01" the first value that is assigned to the event TRIAL_start is the integer 1 at 59453797μs (about 1 minute) after the MWorks server had been started. This information is saved in amm-MSTm-sun-120-01+01-task.h5 as event_value/TRIAL_start[0] = 1 and event_time/TRIAL_start[0] = 59453797 (note that if reading the data into Matlab, the index needs to be 1, rather than 0). In other words, event_time/TRIAL_start is an array whose length equals the number of trials and whose entries specify the start time of each trial. Correspondingly, event_value/TRIAL_start is an array of the values of TRIAL_start at each timestamp. In the case of TRIAL_start, these values are the trial numbers. For events that describe stimulus attributes, such as the event STIM_MappingProbe_posX – which describes the x-coordinate of the random dot pattern that was presented in Experiment 1 (see Methods) – this would be a list of x-positions. A description of events is included in each file in the form of attributes and additionally provided as a separate table in the repository for easy reference. The eye-files are structured in the same way, with event_value/EYE_x_dva, for example, specifying the x-coordinate of the monkeys gaze and event_time/EYE_x_dva specifying the corresponding time stamps, which are synchronized with the time stamps of the corresponding task-file.

4.6 Code availability

We provide MATLAB code that serves two purposes: (a) to give examples of how the data can be accessed; and (b) to perform some elementary data validation analyses. The code is available along with the data in the repository. We briefly summarize the MATLAB scripts that we provide along with the data.

h5_extract

h5_extract() is a function that takes as its input the name of an HDF5 file, a cell array of strings that specify the parameters of the experiment that are to be extracted (typically event_value, and event_time), and a cell

array of strings that specify the events that are to be extracted (typically the list of event names that can be extracted from the HDF5 itself, as demonstrated in the example scripts). The output of the function are two MATLAB structure arrays event_value and event_time which correspond to the groups in the HDF5 file of the same name. Each field name of the structure arrays is the name of an event and the content in the field corresponds to the values and times associated with the event, respectively. The names of stimulus parameters begin with 'STIM_', descriptions of the trial properties begin with 'TRIAL_', and input/ouput variables (such as button presses and rewards) begin with IO_. Once extracted, the data are ready for visualization and analysis. The scripts technical_validation, example_raster_fr, and example_trial_histogram demonstrate the use of the h5_extract() function to read in all events saved in an HDF5 file.

technical_validation

The technical_validation script was used to perform all the plausibility checks described in the section "Technical Validation". The script prints all the statements from that section of the manuscript that contain quantitative information about the data (such as, for example, average firing rate) to the console.

example_raster_fr

To demonstrate how individual spike times can be accessed and visualized, the script creates a raster plot of spike trains and a scatter plot of firing rates per trial (for trials lasting at least 500 ms), colored by trial outcome, for one example file. The example file is specified in the first line of the script and can easily be changed by the user. The raster plot also includes the time of reward delivery and the end of each trial.

example_trial_histogram

To demonstrate how multiple files can be accessed for population analyses, a histogram of the number of trials across recording sessions is produced for each of the three experiments and color-coded by monkey.

example_probe_time

This script contains a function, probe_time(), which extracts the values of all events at a given time point, as well as some additional code that illustrates the use of the function.

example_rc_stim_extraction

This script demonstrates how the direction and speed values of the RC stimulus can be extracted and in particular, how those segments that were masked in some of the files (see Methods) can be determined.

example_eye_data

To demonstrate how the eye tracking data can be accessed, this function plots the x- and y-position of the monkey's gaze as well as the size of the right and left pupil for the first n timesteps (where n can be specified by the user, default is 100).

example_spatial_mapping_analysis and example_tuning_analysis

To demonstrate how spiking activity can be related to stimulus features, these two functions plot firing rate as a function of probe location in the Spatial Mapping experiment (example_spatial_mapping_analysis) or as a function of motion direction and speed in the Tuning experiment (example_tuning_analysis).

Chapter 5

Response properties of MST neurons based on complex motion stimuli

Benedict Wild, Amr Maamoun, Yifan Mayr, Stefan Treue

in preparation

Chapter 2 described MST as an area at the intersection of sensation and cognition. To investigate the sensory aspects of the area in more detail, we recorded the activity of individual MST neurons in response to a variety of different motion stimuli. The experiment was designed by Amr Maamoun, who also collected the majority of the data (63%) during his doctoral studies (Maamoun, 2018).

Author contributions

A.M. and S.T. designed research. A.M. and B.W. performed research. B.W., A.M., and Y.M. analyzed data. B.W. wrote the manuscript. S.T. supervised the project.

Acknowledgements

We thank Leonore Burchardt, Sina Plümer, and Dirk Prüsse for assistance in animal handling, training, and surgeries, the German Primate Center's veterinary and animal husbandry staff for their expert animal care, and Ralf Brockhausen for technical support. This work was supported by grants to

S.T. from the Deutsche Forschungsgemeinschaft (DFG): Collaborative Research Center 889 “Cellular Mechanisms of Sensory Processing” (Project C04) and Research Unit 1847 “Physiology of Distributed Computing Underlying Higher Brain Functions in Non-Human Primates”.

Abstract

Describing the functional relationship between neural activity and the features of external sensory stimuli is one of the fundamental goals of systems neuroscience. For neurons in many visual areas of the primate cortex, such a relationship has been established either by means of tuning functions, or through more complex models of the neural response profile based on cross-correlations between spike trains and a random stimulus. However, few attempts have been made to apply these latter methods to areas beyond early visual cortex. We present three experiments that apply both of these approaches to single neurons in the medial superior temporal cortex (area MST) of the macaque brain. Using classical tuning curve approaches, we confirm that these neurons have large, but well-defined spatial receptive fields and are independently tuned for linear and spiral motion, as well as speed. We also confirm that the tuning for spiral motion is position invariant in a majority of MST neurons. Because measurements with a restricted stimulus set risk missing important aspects of the neuronal response function, we developed a new stimulus that generates smooth, complex motion patterns (nevertheless described by a limited number of parameters), for a bias-free characterization of receptive field profiles. While these profiles are predictive of some of the tuning properties of MST neurons, they are generally less informative than in earlier visual areas and tend to miss features, such as position invariance. Thus, MST emerges as a model system for studying more complex relations of neural activity and external stimuli.

Significance Statement

Understanding how stimuli and behavior are encoded by neural activity is a central goal of systems neuroscience. In the visual domain, this question of neural coding can be phrased as, "How are visual features (such as color or motion) represented in the spiking activity of cortical neurons?". Substantial progress has been made to answer this question for early parts of the visual system, such as the retina or early visual cortex. However, there are only few cases of applying such approaches to more specialized mid- and high-level visual areas. Here we employ a variety of approaches to explore how complex visual motion is encoded by neurons in the medial superior temporal area of the macaque cortex.

5.1 Introduction

Characterizing the relationship between a neurons activity and external stimuli is a central goal of visual neuroscience. Over the past 60 years, two different approaches have emerged to address this challenge: First, stimuli that are parameterized by one or two dimensions (such as orientation, direction of motion, speed, spatial frequency, or color) have been used to describe a neurons "tuning" for these features. This approach was particularly successful in demonstrating tuning for a luminance bars orientation in V1 (Hubel & Wiesel, 1968), for the direction of a moving stimulus in the middle temporal (MT) area (Albright, 1984; Maunsell & Van Essen, 1983c; Snowden et al., 1992), or for color in extrastriate area V4 (Kusunoki, Moutoussis, & Zeki, 2006) of the primate cortex. However, using a highly restricted stimulus set raises the risk of missing out on essential aspects of the neuron's response function. Thus, an alternative approach that makes less assumptions about the stimulus subspace that a neuron responds to has been developed based on triggered correlation (De Boer & Kuyper, 1968). Such "reverse correlation" approaches describe a neurons response to stimuli by a linear filter (Ringach & Shapley, 2004), followed by a non-linearity and a spike generation process, such as a Poisson process (linear-nonlinear Poisson (LNP) models, e.g., Chichilnisky, 2001) or an integrate-and-fire process (Paninski, Pillow, & Simoncelli, 2004; Pillow, Paninski, Uzzell, Simoncelli, & Chichilnisky, 2005). The linear filters have been successfully estimated in cat in primate V1 using the spike-triggered average (STA; e.g., DeAngelis et al., 1993a, 1993b; Jones & Palmer, 1987) or spike-triggered covariance (STC; e.g., Rust et al., 2005; Touryan et al., 2002). However, nearly all of the studies using such approaches recorded spiking activity in the early stages of the visual processing pipeline, such as retinal ganglion cells (RGCs), the lateral geniculate nucleus (LGN) of the thalamus, or primary visual cortex (V1). Mid- and high-level areas of the primate cortex, such as V4 and IT in the ventral pathway and MT and MST in the dorsal pathway, have larger receptive fields, which allows them to integrate over larger parts of the visual field, and respond to more complex aspects of stimuli. This transformation of response functions along the cortical hierarchy has been described as a change from "sensation to perception" (Treue, 2003). Here we used both tuning curves and spike-triggered analyses to characterize the neuronal response function in one higher visual area, the medial superior temporal cortex (MST). Neurons in this area are motion sensitive and direction selective (Tanaka et al., 1986). In contrast to neurons in area MT, which provide the major input to MST, they also respond to more complex motion patterns, such as rotation and radial motion (contraction and expansion) (Duffy & Wurtz, 1991a, 1991b; Saito et

al., 1986). These response preferences for different motion patterns can be integrated in Gaussian tuning curves within a 1-dimensional "spiral motion space" (Graziano et al., 1994). The motion patterns that make up this spiral space are similar to the "optic flow" patterns that are experienced by an observer moving through the environment (Gibson, 1950) and based on their tuning for these optic flow patterns, MST neurons have been suggested to play an important role in self-motion perception (Britten, 2008; Duffy, 1998; Duffy & Wurtz, 1995). This has been supported by the finding that microstimulation of MST neurons systematically biases monkeys heading perception (Britten, 1998). MST, but not MT neurons also show direction selective delay activity in a working memory task (Mendoza-Halliday et al., 2014) and represent illusory flow motion (Luo et al., 2019), suggesting that activity in MST is more closely related to behavior than activity in earlier areas. Thus, MST combines features of early, sensory areas and higher areas involved in perception and cognition (Wild & Treue, 2021b). To explore the response functions of MST neurons in more detail, we conducted three experiments. In the first experiment ("Spatial Mapping") we explored neurons' spatial receptive fields. The second experiment ("Tuning") used the classic tuning curve approach to test how neurons' firing rates depend on linear (translational) and spiral direction as well as speed. In the last experiment ("Reverse Correlation") we recorded neuronal responses to a newly developed motion stimulus that is well suited for spike-triggered analyses. This approach allowed us to estimate the response function of MST neurons in an unbiased manner which can then be used to predict a neuron's response to any arbitrary stimulus. To test this, we used the response functions determined in the Reverse Correlation experiment to predict the responses in the Tuning experiment.

5.2 Methods

Research with non-human primates represents a small but indispensable component of neuroscience research (Buffalo et al., 2019). The scientists in this study are aware and are committed to the great responsibility they have in ensuring the best possible science with the least possible harm to the animals (Roelfsema & Treue, 2014).

Ethics statement

All animal procedures of this study have been approved by the responsible regional government office (Niedersaechsisches Landesamt fuer Verbraucherschutz und Lebensmittelsicherheit [LAVES]) under the permit numbers

3392 42502-04-13/1100 and 33.19-42502-04-18/2823. We have established a comprehensive set of measures to ensure that the severity of our experimental procedures falls into the category of mild to moderate, according to the severity categorization of Annex VIII of the European Unions directive 2010/63/EU on the protection of animals used for scientific purposes (see also (Pfefferle et al., 2018)). The animals were group-housed with other macaque monkeys in facilities of the German Primate Center in Goettingen, Germany, in accordance with all applicable German and European regulations. The facility provides the animals with an enriched environment (including a multitude of toys, wooden structures, and other enrichment; Berger et al., 2018; Calapai et al., 2016), natural as well as artificial light, and a space exceeding the size requirements of the European regulations, including access to outdoor space. The German Primate Center has several staff veterinarians that monitor and examine the animals and consult on procedures. During the study, the animals had unrestricted access to food and fluid, except for the days when data were collected or the animal was trained on the behavioral paradigm. On these days, the animals were allowed unlimited access to fluid through their performance in the behavioral paradigm. Here, the animals received fluid rewards for every correctly performed trial. Throughout the study, the animals psychological and medical welfare was monitored by the veterinarians, the animal facility staff, and the lab's scientists, all specialized in working with nonhuman primates. The three animals were healthy at the conclusion of our study and were used in follow-up studies.

Subjects

Three male rhesus monkeys (*Macaca mulatta*) were tested in the experiments: Monkeys S (age and weight during participation in experiment: 14-16 years; 9.9-12.8 kg), monkey I (10-13 years; 10.4-13.4 kg), and monkey E (14 years; 8.0-9.2 kg). All animals had previously participated in other projects and were implanted with a titanium head holder to minimize head movements during the experiment, as well as with a recording chamber implanted on top of a craniotomy over the parietal lobe based on a magnetic resonance imaging (MRI) scan. Recordings were performed from the left hemispheres of monkeys S and E, as well as from both hemispheres of monkey I. Surgeries were conducted under general anesthesia and post-surgical care using standard techniques.

Task

In all three experiments, monkeys performed a simple eye fixation task to keep their gaze stable throughout stimulus presentation. A red fixation point was presented on the screen and the monkey initiated the stimulus presentation by pressing a button attached to the primate chair. The location of the fixation point was chosen so as to ensure maximal coverage of the spatial receptive field by the screen (typically this meant that the fixation point was placed in the visual hemifield ipsilateral to the recording chamber as most neurons had spatial receptive fields in the contralateral hemifield). The luminance of the fixation point changed within a time window that was randomly selected from a uniform distribution ranging from 3 to 4.6 s after stimulus onset, to which the monkey had to respond within 600 ms in order to receive a fluid reward (juice, tea, or water, depending on each monkeys preferences). Eye positions were sampled binocularly at 500 Hz with an Eyelink 1000 system (SR-Research, Ottawa, ON, Canada). Gaze direction was controlled during the recordings to stay within a square window of 3° side length around the fixation spot. Trials on which the monkeys gaze went outside this window were aborted without reward. Stimuli were presented on a 27" LCD monitor (XL2720T, BenQ, Taipei, Taiwan) or projected onto a 171.5×107.2 cm back projection screen (dnp Black Bead, Karlslunde, Denmark) using a projector (projection design, Fredrikstad, Norway) with a resolution of 1920×1200 pixels and a refresh rate of 60 Hz. All experiments were controlled by the open source software MWorks (`mworks-project.org`) running on an Apple Mac Pro (Apple Inc., Cupertino, CA, USA).

Neural data acquisition

Action potentials ("spikes") were recorded extracellularly with single tungsten electrodes (Thomas Recording, Giessen, Germany) that were advanced using either a hydraulic micropositioner (David Kopf Instruments, Tujunga, CA, USA) or a mechanical microdrive ("Mini Matrix", Thomas Recording, Giessen, Germany). Neural signals were digitized at 40 kHz at 16-bit precision using an Omniplex acquisition system (Plexon, Dallas, TX, USA). The raw signal was filtered with a 6-pole Bessel high pass filter, spike waveforms were detected based on a manually determined threshold, and were then manually sorted into clusters based on different features as implemented in the Plexon OfflineSorter V3 (Plexon, Dallas, TX, USA).

General data analysis

All data analysis was performed in Matlab (The Mathworks, Natick, MA). All analysis steps requiring circular statistics were performed using the CircStat toolbox (Berens, 2009). More details about the analysis steps for each of the three experiments are provided in the following sections.

Experiment 1: Spatial Mapping

Stimuli and procedure

A small, circular random dot pattern (RDP) of radius 1.5° was presented sequentially in different locations across the screen. The RDP consisted of 28 white dots with a radius of 0.2° that were randomly placed and moved independently along linear trajectories in random directions at a speed that was randomly chosen from a range of 4°/s 24°/s. Each presentation of the RDP for 50 ms constitutes one "sample". For every sample, the location of the RDP was updated by assigning a random x- and y-position in the range covering the entire width and height of the display in steps of 2°. On trials that were not interrupted by fixation breaks or incorrect responses (see Task), more than 80 samples could be presented.

Data analysis

The goal of the "Spatial Mapping" experiment was to determine the size and location of a neuron's receptive field. To do so, we extracted the spiking response to every sample, i.e., the number of spikes that a neuron fired during an analysis window of the same length as the sample presentation time (50 ms) that could be shifted in time by a latency period with regard to the sample presentation (Fig. 5.1).For a latency of 75 ms, for example, we would count the spikes that the neuron fired in the time period from 75 ms after sample onset until 125 ms after sample onset (75 ms after sample offset). For each possible sample location, we averaged the number of spikes across repeated presentations in that location and converted the number of spikes to a firing rate in spikes/s (i.e., we multiplied the average number of spikes by 20). We argue that the optimal latency, i.e., the 50 ms time window during which the spiking activity is most likely to reflect a neurons response to the visual input, would lead to the highest variability across the different locations, because at the optimal latency, there should be strong responses to probes presented inside the receptive field and very weak responses to probes presented outside the receptive field. Thus, we calculated the variance in firing rate across all possible probe locations for latencies from 0 to 150 ms in steps of 1 ms and

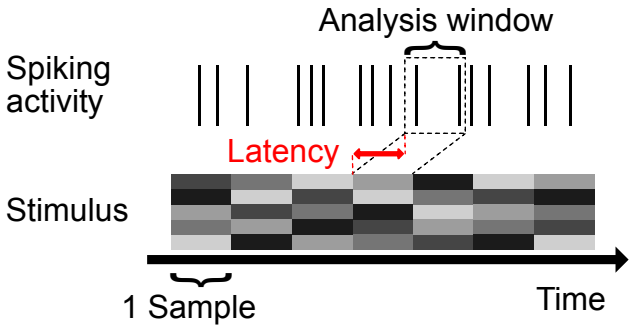

Figure 5.1: Definition of "latency". In each trial, multiple samples of the stimulus are presented. For each sample, the parameters that define the stimulus (e.g., x- and y-coordinate of the mapping stimulus, or x- and y-coordinate, direction, and speed for the tuning stimulus) are updated. The momentary parameters of the stimulus are abstractly symbolized here by each stack of shades of gray. The analysis window has the same length as a sample and is shifted in 1-ms increments in time (the "latency"), relative to the stimulus sample. We consider all spikes that occur during an analysis window as the neuron's response to that sample.

defined the optimal latency as the latency that led to the largest variance (Fig. 5.3A). Visual inspection of the resulting receptive field maps confirmed that this approach identified the latency for which the spatial receptive field was most clearly defined. We then fit a two-dimensional elliptical Gaussian function to the data, using the formula

$$f(x,y) = B + A\exp(-a(x-x_0)^2 + 2b(x-x_0)(y-y_0) + c(y-y_0)^2) \quad (5.1)$$

with

$$a = \frac{cos^2\theta}{2\sigma_x^2} + \frac{sin^2\theta}{(2\sigma_y^2)}$$

$$b = \frac{-sin2\theta}{4\sigma_x^2} + \frac{sin2\theta}{4\sigma_y^2}$$

$$c = \frac{sin^2\theta}{2\sigma_x^2} + \frac{cos^2\theta}{2\sigma_y^2}$$

where B is the baseline firing rate, A is the amplitude of the 2D Gaussian, x_0 and y_0 are the coordinates of the center of the 2D Gaussian, θ is the angle of rotation of the ellipse, and σ_x^2 and σ_y^2 are the x and y spreads of the ellipse. We defined the receptive field size as the area of an ellipse with width $2\sigma_x^2$

and height σ_y^2 and the eccentricity as the distance from the fixation point to the receptive field center defined by x_0 and y_0. Thus, for every cell we extract 3 values from the Spatial Mapping experiment: an optimal latency, a receptive field size, and a receptive field eccentricity. Across the population of all recorded neurons, we explored the distribution of optimal latencies as well as the relation between receptive field eccentricity and size using standard descriptive statistics. We used the coefficient of determination (R^2) to assess the goodness-of-fit and excluded all cells from further analysis that had a R^2 value of less than 0.22.

Experiment 2: Tuning

Stimuli and procedure

A circular RDP was presented sequentially in up to 5 different locations covering the spatial receptive field as it had been determined by an online analysis of Experiment 1. The RDP's radius varied from 5° to 10°, depending on the size of the receptive field and the dot density was 2 dot/deg^2. across the screen. Each presentation of the RDP for 100 ms constitutes one "sample". For every sample, the location of the RDP was updated by assigning an x- and y-position, one of 8 possible linear or spiral directions (from 0° to 315° in steps of 45°), and one of 6 speeds (from 4°/s to 24°/s in steps of 4°/s). The 8 spiral directions were expansion, outward clockwise spiral, clockwise rotation, inward clockwise spiral, contraction, inward counterclockwise spiral, counterclockwise rotation, and outward counterclockwise spiral. For most of the cells, 5 locations were arranged in a cloverleaf, similar to a previous study (Graziano et al., 1994). For cells whose receptive field was on the edge or in the corner of the screen, the arrangement was different to ensure maximal coverage of the receptive field by the 5 RDP positions and in some cells less than 5 locations were probed. Spiral directions were presented in all (up to 5) locations to test for position invariance (see "Data Analysis" below), but linear directions were only presented in the center location. Data analysis The goal of the "Tuning" experiment was to explore MST neurons' tuning for linear and spiral direction of motion, as well as speed. To do so, we extracted the spiking response to every stimulus presentation, i.e., the number of spikes that a neuron fired during an analysis window of the same length as the sample presentation time (100 ms) that could be shifted in time by a latency period with regard to the sample presentation, similar to the "Spatial Mapping" experiment (Fig. 5.1). We argue that at the optimal latency, a well-tuned neuron should show a high variance in its responses to different directions, but a low variance to repeated presentations of the same direction.

In order to determine the optimal latency we calculated the variance across repeated presentations of the same direction for each one of the 8 directions (8 "within-variance" values) as well as the variance across the averaged responses to each of the 8 different directions (1 "across-variance" value) and divided the "across-variance" by the mean "within-variance". Calculating this quotient of across-variance and averaged "within-variance" for every latency from 0 to 150 ms in steps of 1ms, we defined the latency at which this value reached its maximum as the optimal latency. While it is possible that the optimal latency would differ for different locations (e.g., because the neuron receives input from different subpopulations of MT neurons), manual inspection of optimal latencies calculated for each location separately did not show any systematic relation between these two variables. Similarly, there was no systematic effect of speed so that we combined data points from all speeds. For direction tuning we followed the sampling-based, Bayesian approach described by Cronin, Stevenson et al. (2010) for estimating the parameters of the curve and error bars, as well as hypothesis testing. In brief, we estimated the posterior probability distribution of the tuning curve parameters for a circular Gaussian, given the stimulus and the spiking responses. We assumed Poisson spiking to describe the relationship between the response predicted by the tuning curve and the observed data. Markov chain Monte Carlo (MCMC) sampling was used to estimate the posterior parameter distribution. Following the default values suggested by Cronin, Stevenson et al. (2010) in their paper and their toolbox, we us 10,000 burn-in samples, drew 20,000 samples from the posterior and kept only every 50th sample (400 samples) in total to reduce correlations between the samples. These 400 samples were used to correlation 95% confidence intervals. For each cell we fitted up to 42 direction tuning curves based on linear motion presented in one location, spiral motion presented in up to 5 locations, each at six different speeds and one tuning curve that ignored speed. We consider parameter estimates of these fits for our analysis if the following criteria are fulfilled:

- The coefficient of determination (R^2) as a measure of goodness-of-fit is higher than 0.7;
- The 95% confidence interval of the estimated preferred direction covers less than 90° i.e., a quarter of the range of all possible directions;
- The 95% confidence interval of the amplitude does not cross 0;
- The amplitude of the tuning curve is at least 3 spikes/s;
- The maximum firing rate is larger than the baseline firing rate + 10%;

- The maximum firing rate is at least 10 spikes/s.

We considered a cell tuned for linear motion or tuned for spiral motion at one of the tested locations if at least 2 out of the 6 fits (for 6 different speeds) met these criteria. For the cells that survived these criteria, we first tested whether the preferred direction in linear or spiral space depended on speed. We considered a neuron to show speed-dependent direction selectivity (SDDS) if the preferred directions across 6 speeds cover a range of more than 45°. To assess the neurons tuning for speed, we followed the procedure by Churchland et al.(A. K. Churchland et al., 2007) and fit the data with a cubic-smoothing spline using the Matlab function *csaps* with the smoothing parameter set to 0.04. We interpreted the maximum of the fitted curve as a neurons preferred speed. To explore whether responses to spiral motion were position invariant, as had previously been reported (Graziano et al., 1994), we first attempted to reproduce Graziano et al.'s results by following their approach. We included all cells that showed tuned responses to spiral motion in at least 4 locations. Only responses to the preferred and the anti-preferred direction at the center location were considered (in cases where RDP locations were not arranged in a cloverleaf, the location at which linear motion had also been presented was considered the "center location"). Directional selectivity (DS) was calculated as

$$1 - (R_{anti-pref}/R_{pref})$$

where R_{pref} is the response to the preferred direction at the center location and $R_{anti-pref}$ is the response to the direction 180° away from the preferred direction at the center location. For each cell, three or four position invariance indices (PI) were calculated by dividing the DS at the surrounding locations by the DS at the center location. PI values close to 1 indicate position invariance while negative PIs indicate a reversal in preferred direction at that location. To take information from the full tuning curve into consideration, we also calculated the range of preferred directions across all locations for these cells. A position invariant cell should have very similar preferred directions in all locations and therefore a small range, whereas a range close to 180° indicates a reversal in preferred direction in at least one location.

Experiment 3: Reverse Correlation

Stimuli and procedure

A newly developed, large, rectangular random dot pattern ("RC Stimulus") was presented. The stimulus consists of 10×15 square segments, each with a side length of 3°, resulting in an overall size of $30° \times 45°$. For some recordings,

a smaller version of the stimulus with 6×9 segments (overall size: $18° \times 27°$) was used; for the sake of readability, we describe all analysis steps for the 10×15segments case, but all steps can equally be performed for the smaller version of the stimulus. Each segment is assigned a random linear direction from 0° to 360° and a random speed from 0°/s 20°/s and contains 10 dots (1500 dots in total, dot density of 1.11 dots/deg^2). The direction and speed of each dot is determined by Gaussian weighting fields of all segments in its vicinity: A dot that is positioned right in the center of a segment will move in the direction and with the speed assigned to that segment, whereas a dot that is positioned at the intersection of four segments will move in a direction and with a speed that is the average of the values assigned to these four segments. The influence of each segment on the dots in its vicinity is defined as a 2D-Gaussian that has its mean at the center of the segment and a standard deviation of 1.2°. The direction and speed values assigned to each segment are updated every 100 ms, the duration that constitutes one "sample". Overall, what the observer perceives is a smooth, wave-like pattern of moving dots. On trials that were not interrupted by fixation breaks or incorrect responses (see Task), more than 40 samples could be presented.

Data analysis

In order to reduce the number of parameters that need to be fitted, we ignored the different speeds in all analyses presented here and represent each sample of the RC Stimulus as a vector of 150 directions (Fig. 5.6A). The stimulus for an entire recordings session is then a number-of-samples $\times$150 matrix. Analysis thus becomes a regression problem where we try to predict the spike count from the stimulus vector. Because regression analysis is problematic with circular data, we converted the 150 directions to 300 x- and y-components by taking the cosine and sine of each direction. The final version of the research problem is thus to predict the [number-of-samples $\times$1] spike counts vector (y) from a [number-of-samples $\times$300] design matrix (X). We test five models of increasing complexity. To assess the performance of each model, we simulated the stimuli used in the "Tuning" experiment (see above) by calculating for each segment of the RC Stimulus the local direction that the "Tuning" RDP would have had in that location (Fig. 5.6B). For recordings on which the "Tuning" RDP had been presented in 5 locations, we would thus generate 48 such stimuli: 8 spiral directions in 5 locations plus 8 linear directions in one location. To quantify each models ability to predict a neurons response to the "Tuning" stimuli, we calculated four measures: (1) the coefficient of

determination (R^2) of the model, calculated as

$$1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (y_i - \bar{y})^2}$$

where $\hat{y}_i$ is the predicted response for sample i, y_i is the measured response for sample i and $\bar{y}$ is the mean across all measured responses; (2) the correlation between the predicted and the real responses from the "Tuning" experiment (using the responses from the "Tuning" experiment that were calculated across all speeds); (3) the difference in preferred direction between tuning curves fitted to the predicted responses and tuning curves fitted to the real responses; and (4) the difference in tuning width (defined as 1 standard deviation of a Gaussian tuning curve) between tuning curves fitted to the predicted responses and those fitted to the real responses. The first model we tested is a General Linear Model with Gaussian noise, i.e., standard linear regression. We added a column of 1s to the design matrix as an offset parameter and calculated a whitened spike-triggered average (STA) by multiplying the inverted covariance matrix of the design matrix with the normal STA (design matrix multiplied with response vector):

$$(X^T X)^{-1} X^T y$$

The second model was a Poisson Generalized Linear Model (GLM) with an exponential nonlinearity that was fitted using Matlab's *glmfit* function. The third model differed from the second only in that we replaced the default exponential nonlinearity with a nonparametric estimate by binning the output of the linear filter, computing the average number of spikes for each bin and using Matlabs *interp1* function to extend these individual points to a full function that can be evaluate at any arbitrary point. The fourth model was a variant of the second model in which we added ridge regularization by adding a penalty on the sum of squared regression coefficients that was multiplied with a ridge parameter λ. We chose the value of λ by splitting the data for each recording in a training (80%) and a test (20%) set, fitting the model with varying values of λ to the training data, and selecting the value that had the lowest error on the test data. For the fifth and final model we employed the information-theoretic generalization of spike-triggered average and covariance analysis (iSTAC) suggested by Pillow & Simoncelli (2006). Using the code they provide, we calculated the STA and spike-triggered covariance (STC), three "iSTAC filters", i.e., orthogonal basis vectors that are sorted in order of their informativeness, and fitted an exponentiated-quadratic nonlinearity using direct maximum-likelihood estimation of the quadratic parameters. This nonlinearity can then convert the convolution of the three filters with the stimulus into a firing rate. To determine the optimal latency for the Reverse

Correlation experiment, we extracted spike counts for latencies from 0 to 200 ms in steps of 5 ms, fitted the first model (linear Gaussian) to each spike count vector, and defined the latency for which R^2 was largest as the optimal latency. All cells for which responses to at least 5000 samples were recorded and that met the inclusion criteria for "Spatial Mapping" and "Tuning" (see above) were included in the analysis of the "Reverse Correlation" experiment.

Data availability

The fully annotated raw data underlying the analyses in this manuscript are publicly available (Wild et al., 2021) and detailed information on how to access and analyze the data is provided in an accompanying manuscript (Chapter 4).

5.3 Results

The aim of our study was to elucidate the functional relationship between neural activity across neurons in area MST and the complex moving stimuli these neurons respond to. We characterized MST single neuron response by conducting three experiments to focus on

- the spatial profile of the receptive fields in MST,
- the neurons tuning to complex motion patterns and
- to assess the responsivity and selectivity for a novel class of multi-dimensional motion patterns through reverse correlation.

We analyzed 171 cells that had been recorded in 138 recording sessions from 4 hemispheres of 3 monkeys. Out of these 171 cells, 29 were excluded from all analyses, because they did not provide interpretable data in any of the 3 experiments (i.e., they did not have a clear spatial receptive field, showed no tuning for linear or spiral direction, and did not respond selectively to the RC stimulus). Table 5.1 shows the distribution of cells across the 4 hemispheres. Recordings that were not included (i.e., the difference between rows 2 and 3 in Table 5.1) were either too noisy for spike sorting or cells were lost before a sufficient number of trials could be collected

General neural activity statistics

Aspects of neuronal activity, such as baseline firing rate, maximum firing rate, or firing rate variability, can vary across different areas of the primate brain.

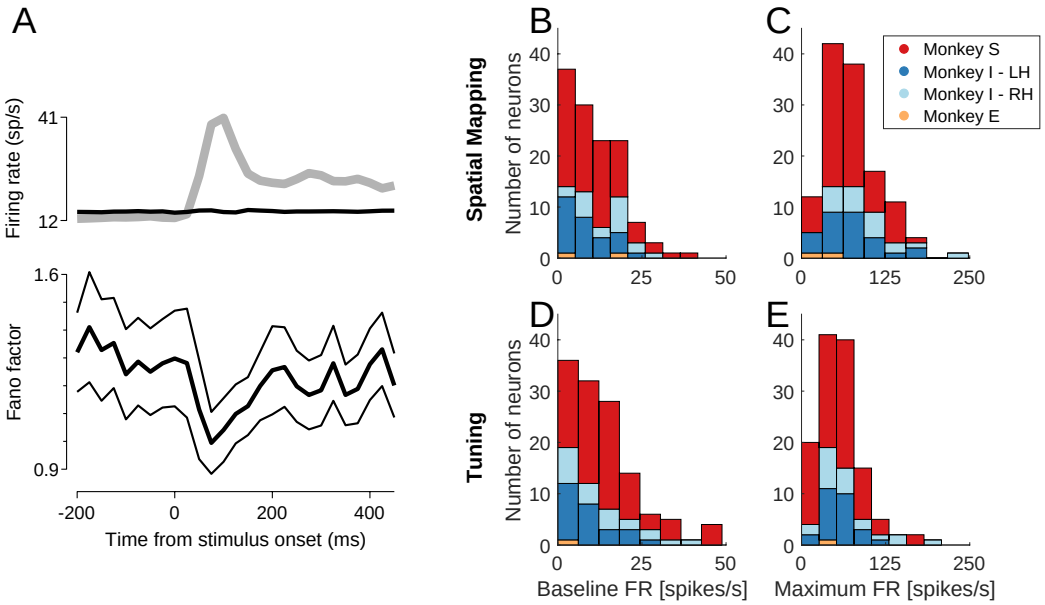

Figure 5.2: General statistics of neural activity. ***A***. Changes in mean firing rate and firing rate variability in the "Tuning" experiment after stimulus onset. The mean rate (top panel in gray) and the Fano factor (bottom panel, black with flanking standard error) were computed using a 50 ms sliding window. The Fano factor was computed after mean matching and the resulting stabilized means are shown in black (see Churchland et al. (2010) for details). Note the similarity to the results for other brain areas shown in Churchland et al. (2010). ***B***. Distribution of baseline firing rates in the "Spatial Mapping" experiment. ***C***. Distribution of maximum firing rates in the "Spatial Mapping" experiment. ***D***. Distribution of baseline firing rates in the "Tuning" experiment. E. Distribution of maximum firing rates in the "Tuning" experiment. See main text for details. Color of the bars in panels B-E indicates the monkey and hemisphere.

Before we describe the results of our three experiments, we provide general statistics about neuronal activity measures in our recordings of MST neurons. Churchland and colleagues (2010) showed that across many cortical areas of the primate brain such as V1, V4, MT, the lateral intra-parietal areas (LIP), dorsal premotor cortex (PMd), the parietal reach region (PRR), and orbitofrontal cortex (OFC) firing rate variability, as measured via the Fano Factor, decreased after stimulus onset. We applied their analysis code[1] to our "Tuning" data, which is most comparable to the kind of data they analyzed, and confirmed that this decrease in neural variability can also be observed in MST, resulting in a plot that is noticeably similar to the plots in Fig. 3 of Churchland et al. (2010) (Fig. 5.2A). We also show distributions of the baseline and maximum firing rate in both the "Spatial Mapping" and the "Tuning" experiment (Fig. 5.2B-E). A paired-sample t-test for the 113 cells that were included in both, the Spatial Mapping and the "Tuning" analysis

[1] https://churchland.zuckermaninstitute.columbia.edu/content/code

	Monkey S LH	**Monkey I LH**	**Monkey I RH**	**Monkey E LH**	**Total**
Recording attempts	225	126	77	18	446
Recorded files	193	36	29	5	263
Included files	82	26	28	2	138
Cells	108	26	35	2	171
Included in analysis	89	21	30	2	142
Included in Spatial Mapping analysis	76	19	28	2	125
Included in Tuning analysis	78	21	26	1	126
Included in Reverse Correlation analysis	55	9	22	0	86

Table 5.1: Distribution of cells across the 4 hemispheres of three monkeys. LH: left hemisphere; RH: right hemisphere.

showed that the baseline firing rate was significantly higher for the "Tuning" experiment ($t(112) = 2.96, p < .01$), which can be explained by the difference in how the values were calculated (see "Methods"). Importantly, however, both approaches show that some MST neurons show quite high levels of baseline activity, with 9 out of 113 cells showing a baseline firing rate larger than 20 spikes/s in both experiments. Similarly, the maximum firing rate was significantly higher in the "Tuning" than in the "Spatial Mapping" experiment ($t(112) = 6.77, p > 0.001$), which is not surprising, given that larger stimuli with coherent motion were presented inside the receptive field. Seven out of 113 cells reached a maximum firing rate of more than 100 spikes/s in both experiments and 10 cells had a firing rate larger than 100 spikes/s in the tuning experiment.

Spatial Mapping

In the "Spatial mapping" experiment, a small random dot probe was presented sequentially for 50 ms at different locations across the visual field. We averaged the spiking response at each location to estimate the spatial receptive field. By repeating this procedure for a range of latencies between stimulus onset and analysis window and comparing the variance in firing rates across different locations, we determined the optimal latency at which the neuron's activity best reflected its response to the stimulus. We estimated spatial receptive field size and eccentricity by fitting a 2D Gaussian to the data (see "Methods" for details). 17 cells were excluded from the analysis because they did not have a clearly defined spatial receptive field, leaving 126 cells to be included in the analyses (Table 5.1). Fig. 5.3A illustrates that the variance in firing rates across the different locations was indeed a good measure to determine the optimal latency between stimulus and analysis window. Plotting variance as a function of latency (Fig. 3A, left panel) shows an increase in variance with latency up to a maximum at a latency of 59 ms for this cell before decreasing again. Plotting the firing rates as a function of location confirms that the spatial receptive field is much more clearly defined at this optimal latency (Fig. 5.3A top right panel) as compared to a randomly chosen control latency (Fig. 5.3A bottom right panel). The distribution of optimal latencies across all cells shows a wide range, from 12 ms to 102 ms (Fig. 5.3B). An independent t-test showed that the average optimal latency differed significantly between Monkey I (77.9 ms, pooled across both hemispheres) and Monkey S (54.7 ms), $t(122) = 8.95, p < 0.01, d = 1.66$); we did not perform any statistics on the data from monkey E, as there were only 2 cells included. Fig. 3C shows a 2D Gaussian fitted to the data from the same example cell as in Fig. 5.3A. The receptive field size was defined as the area of an ellipse with width and

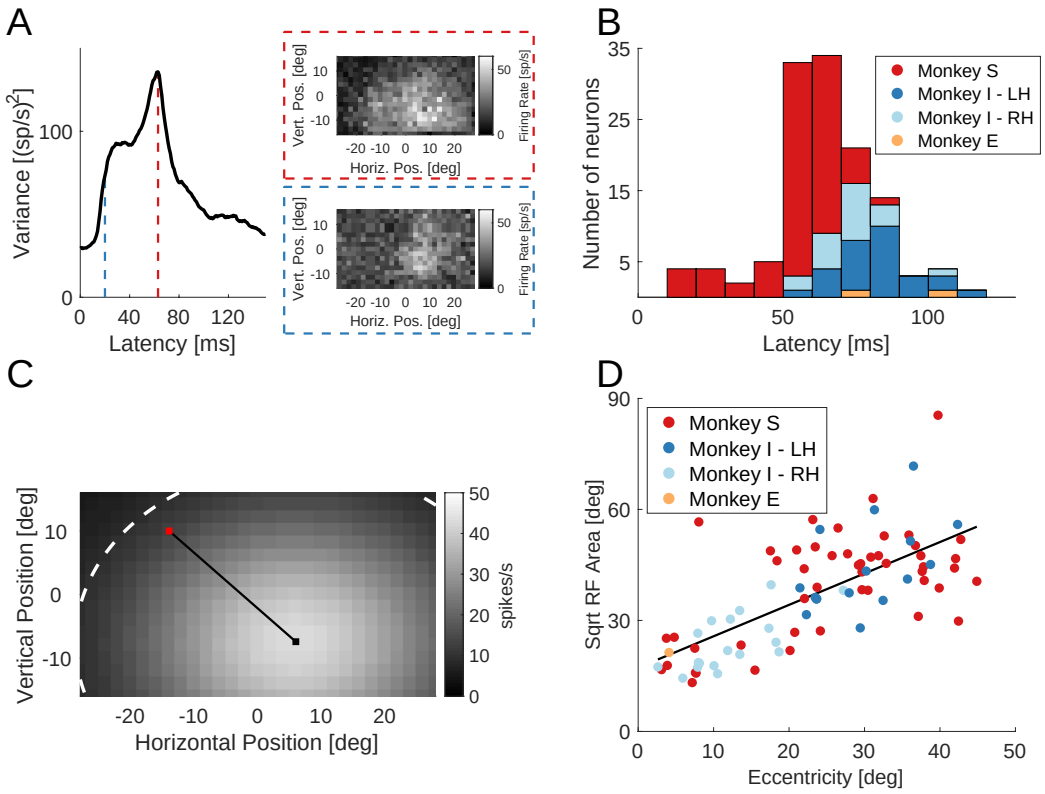

Figure 5.3: Results from the "Spatial Mapping" Experiment. ***A***. Definition of the optimal latency. Left panel shows the variance in firing rate across locations plotted as a function of latency. The two right panels show the receptive field map at the optimal latency (top, red box) and at a randomly chosen control latency (bottom, blue box). The two latencies are also marked in the left panel (see main text for details). ***B***. Distribution of optimal latencies across all cells. Color of the bars indicates the monkey and hemisphere. ***C***. 2D Gaussian fitted to the example cell that was shown in panel A. The white dashed line shows the receptive field outline, defined as an ellipse with height and width set to 2 times the standard deviations of the Gaussian. The red dot at location [-15,10] marks the fixation point. The black dot marks the receptive field center. The line connecting the fixation dot and receptive field center is the eccentricity. ***D***. Receptive field size (square-root of the area) as a function of eccentricity across all cells. The black line shows the linear regression line.

height set to two times the standard deviations of the 2D Gaussian (outlined with the white dashed line). Eccentricity (marked with a black line) was defined as the distance from the fixation point (marked in red) to the center of the 2D Gaussian (marked in black). Before exploring the relationship between receptive field size and eccentricity across the population of neurons, we excluded an additional 41 cells whose receptive field center lay more than 2 degrees outside the screen. While the fitting procedure might still be able to find parameters that describe the data well for these cells, the size and eccentricity estimates provided by the 2D Gaussian fit becomes very unreliable when only a small part of the receptive field is covered by the screen. For the remaining 85 cells we found a highly significant correlation between eccentricity and receptive field size (quantified by the square-root of the area), $r = .67, p < .01$. The slope of a linear regression line was 0.85 and the intercept was 17.23.

Tuning

In the "Tuning" experiment, a random dot probe was presented sequentially for 100 ms at up to 5 different locations covering the receptive field. For each presentation, one of 8 possible spiral motion directions and one of 6 possible speeds was chosen. Additionally, for the location covering the center of the receptive field, one of 8 possible linear directions could be presented. Fig. 5.4A and 5.4C show the ratio of the variance in the responses to different directions to the variance in the responses to the same direction for different latencies (see "Methods" for details). The clear peak, which we use to define the optimal latency, suggests that our measure is valid. Among all cells that were tuned for linear motion ($N = 100$), the average optimal latency for responses to linear motion was 80.3 ± 2.8 ms (mean $\pm$ SEM, Fig. 5.4B). Among all cells that were tuned for spiral motion ($N = 122$), the average optimal latency for responses to spiral motion was 91.2 ± 2.6 ms (Fig. 5.4D). Across the population of cells that were tuned to both, linear and spiral motion, and included in the "Spatial Mapping" analysis (N=84) the optimal latency for spiral motion was significantly larger than the optimal latency for linear motion (paired t-test, $t(83) = 2.62, p < .05$) and optimal latencies were larger for both linear and spiral tuning than for spatial mapping (linear tuning vs. mapping: $t(83) = 5.35, p < .001$; spiral tuning vs. mapping: $t(83) = 8.59, p < .001$).

We fitted tuning curves for direction and speed to extract preferred parameters and tuning strength (see "Methods" for details). Out of 5901 direction tuning curves we excluded 3385 curves (57.4%) from further analysis based on the criteria outlined in "Methods": 2924 curves (49.6%) had an

insufficient fit quality, 321 curves (5.4%) had a confidence interval for their preferred direction larger than 90°, 0 curves (0%) had a confidence interval for their amplitude that crossed 0, 812 curves (13.8%) had an amplitude of less than 3 spikes/s, 741 curves (12.6%) had a maximum firing rate of less than 10% above their baseline firing rate, and 888 (15.1%) had a maximum firing rate of less than 10 spikes/s. Note that many curves met multiple criteria which is why the numbers add up to more than 3385. This led to 17 cells being completely excluded from the analysis as we found them to be tuned neither for linear nor for spiral motion (Table 5.1). Of the remaining 126 cells, 4 were tuned only for linear motion, 26 were tuned only for spiral motion, and 96 were tuned for both, linear and spiral motion. Figure 5.4E and G shows an example of tuning curves for linear and spiral motion, respectively. We first tested whether a neuron's preferred direction varied with speed (speed-dependent direction selectivity, SDDS). Only 23 out of 100 cells (23%) tuned for linear motion showed SDDS. 60 out of 122 cells (49%) tuned for spiral motion showed SDDS in at least one location, but only 23 cells (19%) showed SDDS in more than 1 location and only 8 cells (7%) showed SDDS in more than 2 locations. Overall, these results suggest that MST neurons' preferred direction does not change with speed. We will therefore conduct our further analyses based on the preferred direction that was calculated across all speeds.

The distribution of preferred linear directions across all cells tuned for linear motion did not differ significantly from a uniform distribution ($p = 0.57$, Hodges-Ajne test for non-uniformity of circular data, Fig. 5.4C). For spiral motion, we tested the distribution of preferred directions at the location where linear motion had also been presented for those neurons that were tuned for spiral motion in that location (63 neurons). This distribution also did not differ significantly from a uniform distribution ($p = 0.68$, Hodges-Ajne test, Fig. 5.4F). However, when including the preferred directions from all locations that at which significant tuning curves had been measured, (280 preferred Directions from 95 neurons), the distribution did differ significantly from a uniform distribution ($p < 0.001$, Hodges-Ajne test) with a bias towards expansion and contraction, similar to previous reports (Graziano et al., 1994). With regard to speed tuning, we focused on tuning curves that were fitted across directions as visual inspection suggested that the effect of direction on speed tuning was a change in gain but did not change speed preferences (similar to how speed affected the amplitude, but not preferred direction of direction tuning curves). We found that most cells preferred either the slowest (4 °/s: 28% for linear motion and 29% for spiral motion) or the fastest speed (24 °/s: 56% for linear motion and 50% for spiral motion) that was presented. Only a minority of cells showed a tuning curve that had its

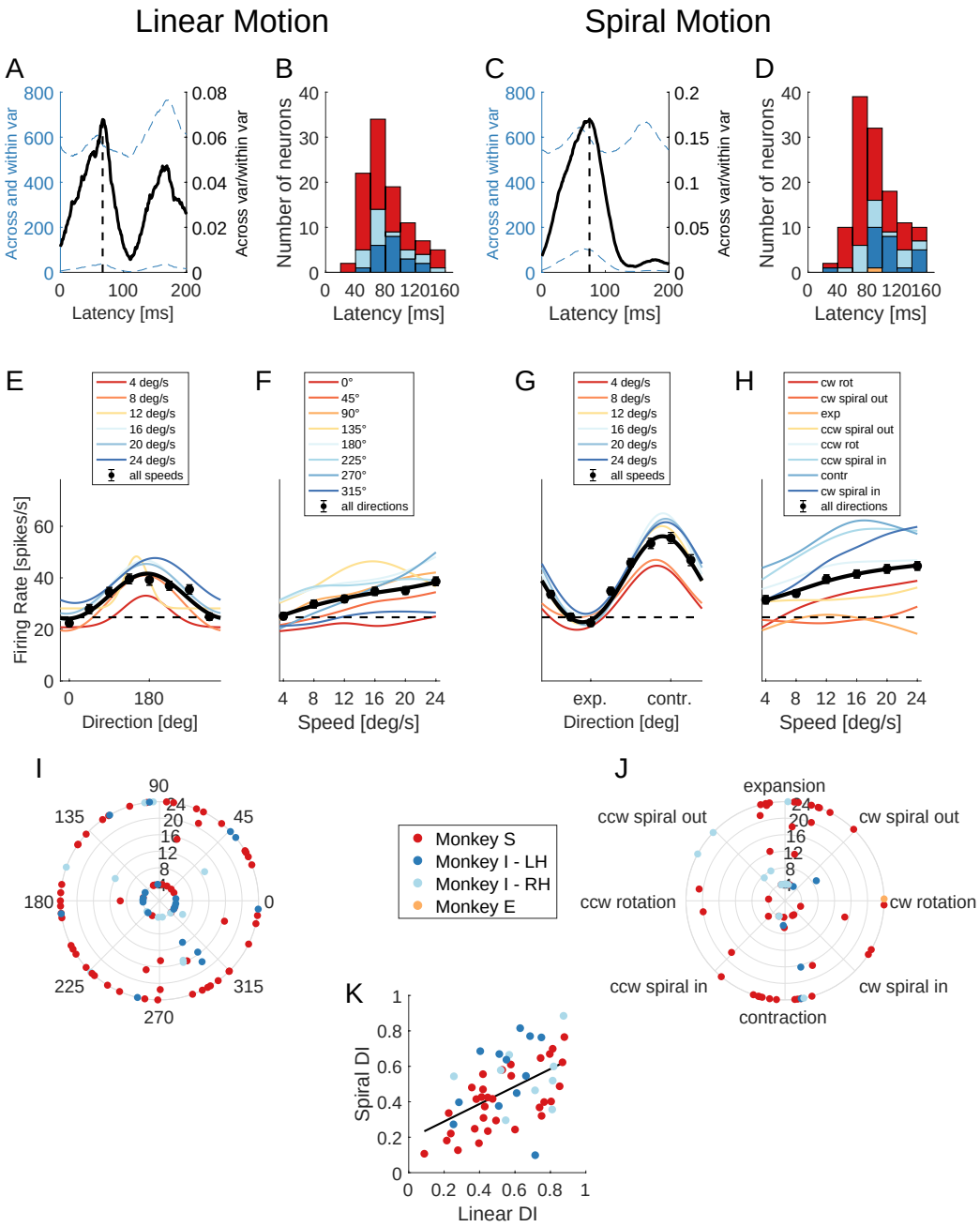

Figure 5.4: Results from the Tuning Experiment. Panels A.-B., E.-F., and I. show results for linear tuning, panels C-D, G-H, and J show results for spiral tuning.

peak somewhere within this range (16% for linear motion and 21% for spiral motion). Next, we looked at cells that were tuned for both, spiral and linear

Figure 5.4: *(continued)* ***A***. Optimal latency for linear motion in one example cell was determined by calculating the variance in spiking responses for repeated presentations of the same direction ("within variance") and across the 8 different directions ("across variance") for every latency (gray dashed lines, left y-axis). Optimal latency (indicated by dashed line) was defined as the maximum of "across variance" divided by "within variance" (solid black line, right y-axis) (see "Methods" for details). ***B***. Distribution of optimal latencies for tuning to linear motion across all cells. Color of the bars indicates the monkey and hemisphere. ***C***. Like A, but for spiral motion. ***D***. Like B, but for spiral motion. ***E***. Example direction tuning curve for linear motion. Colors indicate speed. Black points and line show data across all speeds. Error-bars show standard error of mean. ***F***. Example speed tuning curve for linear motion. Colors indicate direction. Black points and line show data across all direction. Error-bars show s.e.m. ***G***. Like E., but for spiral motion. ***H***. Like F., but for spiral motion. ***I***. Distribution of preferred directions (angle) and speeds (radius) across all cells. Dot color indicates the monkey and hemisphere. ***J***. Like I., but for spiral motion. ***K***. Scatter plot of linear and spiral directionality index. Dot color indicates the monkey and hemisphere. Black line is a linear regression (slope: 0.52).

motion, specifically on those that were tuned for linear motion and for spiral motion at the location at which linear motion had been presented (49 cells). We found that in these cells the directionality indices for linear and spiral tuning were significantly correlated ($r = .60, p < .001$), indicating that cells that are strongly tuned in one motion space are also strongly tuned in the other space. However, the circular correlation between the preferred linear and spiral directions was not significant ($r = .10, p = .48$).

We assessed position invariance for cells that were tuned across multiple locations (see Methods for details). Fig. 5A shows tuning curves for one example cell that had been presented with spiral motion in five different locations covering the receptive field (Fig. 5B). As can be clearly seen, the tuning curves are very similar across the different locations and speeds that were tested, with the preferred directions (averaged across speeds) for the 5 locations ranging from 233 to 273. A distribution of Position Invariance Indices (PIs) was clearly centered on 1 (the value that indicates position invariance) and the mean PI was 0.94. A one-sample t-test indicated that this value was significantly different from 1 ($t(200) = -2.57, p < 0.05$), but note that the PI does not need to be 1 to indicate position invariance: a reversal in preferred direction would actually lead to negative numbers (which was the case for only 1 PI out of 206). A small decrease in PI can simply

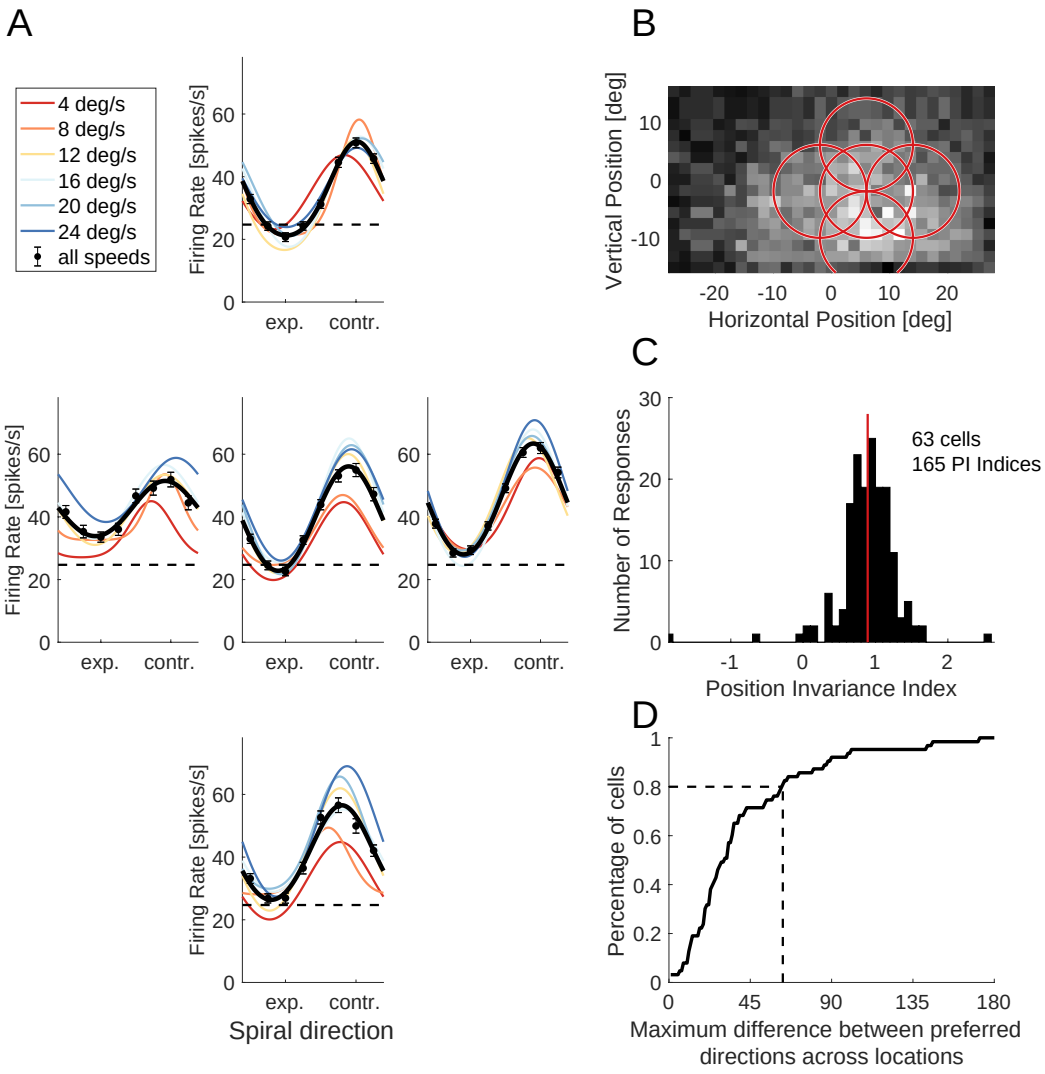

Figure 5.5: Results for position invariance tests. ***A***. Tuning curves at 5 locations for one example cell. ***B***. Stimulus positions for which responses were recorded, plotted on top of the receptive field map. ***C***. Distribution of position invariance (PI) indices. Values close to 1 indicate perfect position invariance. ***D***. Cumulative distribution of the range of preferred directions across locations. For 80% of cells (indicated by horizontal dashed line) the preferred direction across locations fell within an interval of 62 (indicated by vertical dashed line).

be explained by weaker tuning as stimuli are placed towards the edge of the RF. Lastly, we calculated the range of preferred directions for all cells that showed significant tuning for spiral motion in at least 3 locations. As shown in Fig. 5D, 80% of cells had preferred direction within a relatively narrow window of 62°. A cell that would reverse its direction preference would have a range of 180°.

Reverse Correlation

In the "Reverse Correlation" experiment, a newly developed random dot stimulus was presented, which consists of a grid of direction and speed seeds that determine the motion of dots in the vicinity of each seed (Fig. 5.6A). Each seed is assigned a new random speed and direction every 100 ms so that the stimulus at any time can be described as a vector of direction and speed values. In the analyses presented here we ignore speed and aim to predict spiking responses from the vector of directions that define the stimulus at any point in time.

We tested 5 regression models and assessed the performance of each model using 4 metrics: the proportion of variance in the data explained by each model (R^2), the correlation between the responses recorded in the "Tuning" experiment and the responses predicted by the model, the difference in the preferred direction of tuning curves fitted to the real and to the predicted data, and the difference in the tuning curve width. Fig. 5.6C shows a scatter plot of the 48 recorded responses to 8 linear directions and 8 spiral directions in 5 locations and responses predicted by the 4th model (Poisson GLM with regularization) for one example cell. Fig. 5.6D shows the tuning curves fitted to the real data (black) and the simulated data (red). As in this example, all five models generally predicted firing rates that were overall lower and also had a lower amplitude than the actual firing rates. This is likely due to the fact that the models were trained to predict the spike-count in response to the RC stimulus, which is less coherent than the random dot patterns used in the Tuning experiment and therefore generally elicits lower spike counts. For the example cell shown in Fig. 5.6, for example, only 6 out of 19743 samples (0.03%) elicited more than 8 spikes in 100 ms. In contrast, during the "Tuning" experiment 36 out of 1455 presentations of linear motion (2.5%) and 422 out of 7421 presentations of spiral motion (5.7%) elicited more than 8 spikes in 100ms. Apart from the changes in the absolute magnitude of the response, Fig. 5.6D shows that the model follows the data quite well in some cases (Spiral, Location 5), but diverges quite a bit in other cases (e.g., Spiral, Location 4). The summary for each models performance in the four different performance metrics we used is shown in Fig. 5.6E-H. Repeated

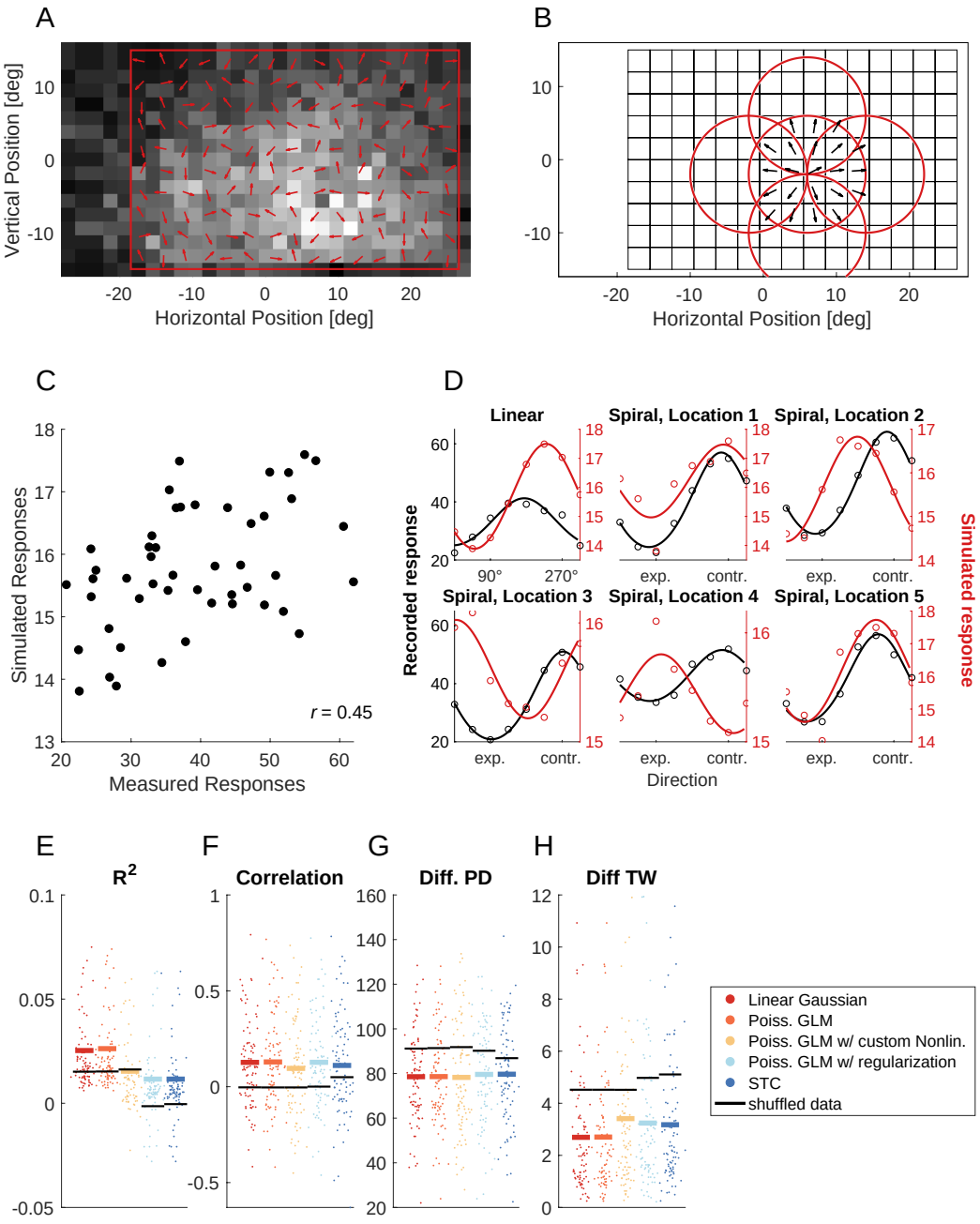

Figure 5.6: . Results of the "Reverse Correlation" experiment. ***A***. Illustration of the "RC stimulus", defined by a 10×15 grid of independently and randomly chosen directions, overlaid with the receptive field map from the "Spatial Mapping" experiment (cf. Fig. 5.3).

Figure 5.6: *(continued)* ***B.*** Example for how expanding motion (direction 90° in spiral space) at the central location from the "Tuning Experiment" was simulated in the framework of the RC stimulus. The grid in the background shows the positioning of the RC stimulus, the red circles show the locations that had been probed during the "Tuning" experiment (cf. Fig. 5.5B), and the arrows indicate the motion assigned to each segment to simulate the "Tuning" motion pattern. ***C.*** Responses recorded during the "Tuning" experiment (x-axis) plotted against responses predicted by Model 2 (Poisson GLM). ***D.*** Tuning curves from the "Tuning" experiment in black and tuning curves fitted to the simulated responses in red. Note the difference in scale for real and simulated responses in panels C and D. ***E.-H.*** Performance of the 5 models (color coded) in 4 different metrics. Dots show individual cells; horizontal, colored bars show the mean; horizontal black lines show the mean of 1000 samples in which the matching of stimulus samples and spiking responses were shuffled. ***E.*** R^2 of the model fit to the data; ***F.*** correlation between real and simulated responses (as illustrated by the example in panel C); ***G.*** mean difference in preferred direction; ***H.*** mean difference in tuning width across all tuning curves.

measures ANOVAs indicate that the 5 models differed significantly in their R^2 values ($F(4, 340) = 60.78, p < .01$) and in the difference between tuning widths ($F(4, 340) = 31.87, p < .01$), but not in the correlation between real and simulated responses ($F(4, 340) = 1.40, p = .23$) or in the differences between preferred directions ($F(4, 340) = 0.20, p = .94$). Surprisingly, the Poisson GLM (model 2) with a standard exponential nonlinearity and without any regularization appeared to perform best, with the highest average $R2$ value and the lowest average difference in tuning width. Generally, R2 and correlation values were relatively low across all models tested, with some outliers reaching R2 values above 0.05 and correlation coefficients above 0.5. We therefore asked whether there was anything that set these neurons, for which the models were able to predict responses well, apart from the rest of the population. We found no significant correlation between either R2 or the simulated/real response correlation of model 2 with the number of samples recorded, indicating that better performance could not be explained by more data. There were, however, significant correlations between model 2s R2 values and both, the linear directionality index (DI) ($r = 0.42, p < .001$) and the spiral DI ($r = 0.25, p < .05$), as well as between the correlation of simulated and real responses and linear DI ($r = 0.38, p < .001$) and spiral DI ($r = 0.30, p < .01$), suggesting that responses of cells that are more strongly tuned can be better predicted.

5.4 Discussion

We present a series of three experiments that explore the response properties of neurons in macaque extrastriate area MST. In particular, our third experiment used a novel random motion stimulus that allows to apply the type of regression analyses that have been very successful in describing early visual areas (e.g., Pillow et al., 2008; Pillow & Simoncelli, 2006; Rust et al., 2005; Touryan et al., 2002) to a high-level visual area, such as MST.

MST combines attributes of low- and high-level visual areas

Our first two experiments ("Spatial Mapping" and "Tuning") replicate findings that MST neurons share properties of typical sensory areas, such as the relation of receptive field eccentricity and area (Fig. 5.3D; see also Desimone & Ungerleider, 1986; Komatsu & Wurtz, 1988b) or the tuning for simple stimulus features, such as direction and speed (Fig. 5.4). However, we also found properties such as position invariance (Fig. 5.5; see also Graziano et al., 1994), which suggest that the representation of motion stimuli in MST is not dominated by low-level physical stimulus features, but more global and abstract.

An interesting finding are the large latencies of close to 100 ms in both the "Spatial Mapping" and "Tuning" experiment (Figures 5.3 and 5.4). Event-related EEG potentials in the human brain reveal activity related to image recognition as early as 150 ms after presentation of a 20 ms stimulus (Thorpe et al., 1996) and reaction times for a simple discrimination task can be as low as 120 ms (Kirchner & Thorpe, 2006). In a simple discrimination task that presumably relies on MST activity ("expansion vs. contraction" or "clockwise vs. anticlockwise rotation") with a low-coherence RDP, subjects responded manually with reaction times of around 600 ms (Strong et al., 2017), but could do so presumably much faster for higher coherence. It seems unlikely then, that 100 ms should be needed simply for the visual signal to reach MST. However, our measure of optimal latency attempts to determine the time window for which the response maximally differentiates between preferred and non-preferred stimuli. This does not mean that basic information about the stimulus cannot be processed much earlier. And indeed, as can be seen in the example in Fig. 5.3A, even though the spatial receptive field is much more clearly defined at the optimal latency (red box), its outline is already visible at an earlier latency (blue box). The fact that the latency at which the response best differentiated between stimuli was significantly delayed for spiral motion, as compared to linear motion, suggests that additional

processing steps are necessary to optimally represent these more complex spiral motion patterns.

One surprising finding in our experiment was the correlation between tuning strength for linear and spiral tuning (Fig. 5.4K). One could well imagine that a cell tuned for linear motion might appear to be tuned for spiral motion as well (and vice versa), because the linear motion resembles local aspects of the spiral motion pattern (Fig. 5.7). In that case, one would expect a negative correlation for tuning strength (because it is not real tuning, but rather spurious tuning based on similarity). Instead, we find the opposite: a strong correlation in tuning strength with no correlation in preferred direction, suggesting that these cells are independently tuned in both, spiral and linear space. This has interesting consequences for decoding (by a human observer or by downstream areas in the brain): how would the decoder know whether a response was elicited by spiral or linear motion? The lack of correlation in preferred direction can actually be helpful here. Any decoding by downstream areas in the brain would rely not on single unit responses, but on a population of neurons with diverse direction preferences. In fact, many computational decoding approaches in MST rely on population activity (e.g., Ben Hamed et al., 2003; Gu et al., 2010; Lappe et al., 1996). From a population of strongly tuned neurons with decorrelated preferred directions in linear and spiral space, one can easily decode whether linear or spiral motion was presented by looking at the correlations of those neurons that share similar preferred directions in one motion space, but not in the other.

Spike-triggered analysis

Quantitative characterization of visual receptive fields using random stimuli and spike-triggered analyses has a long history in systems neuroscience. The spike-triggered average (STA) has successfully been used in the salamander retina (Meister, Pine, & Baylor, 1994), in the primate LGN (Reid & Shapley, 1992), and in simple cells of cat primary visual cortex (DeAngelis et al., 1993a, 1993b; Jones & Palmer, 1987). Spike-triggered covariance (STC) has been used to extract multiple filters and nonlinear combination functions to characterize more complex receptive fields, such as complex cells in cat and monkey V1 (Rust et al., 2005; Touryan et al., 2002). However, few attempts have been made to apply these methods to extrastriate visual areas of the primate cortex. Richert et al. (2013) calculated the STA for responses to a stimulus with 300 independently moving dots that changed direction every 200 ms to characterize receptive fields in area MT. They found, contrary to the classical assumption of a single preferred direction in a single region of space, that many MT neurons had multiple preferred directions, depending on

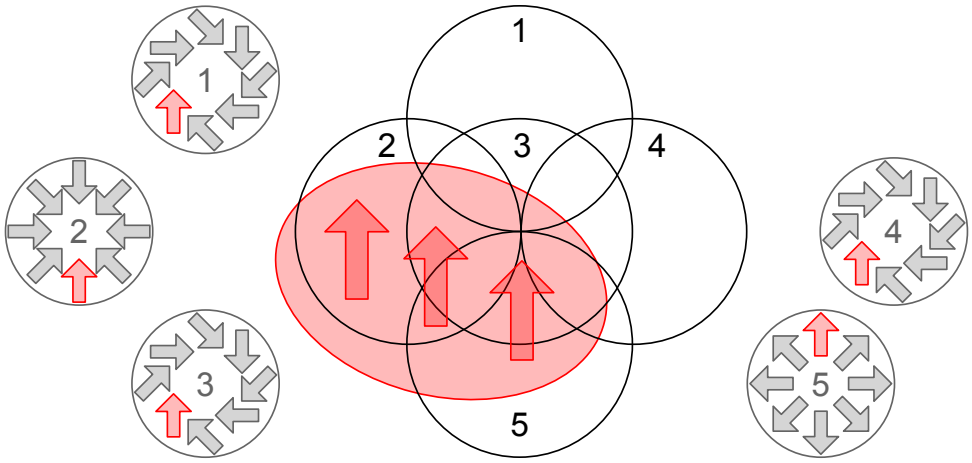

Figure 5.7: Illustration of how a cell that is tuned for linear motion might appear to be tuned for spiral motion. The receptive field of a hypothetical neuron that is only tuned for linear motion (here preferring upward motion) is shown in red. Spiral stimuli are presented in 5 locations (black circles, numbered 1-5). Because of the positioning of the stimuli with regard to the receptive field, stimuli in all 5 locations will evoke responses that appear to show spiral tuning. The gray circles show for each of the 5 locations which spiral pattern has optimally placed local upward relative to the receptive field (red arrows), evoking the strongest response. In locations 1, 3, and 4 only the bottom left part of the stimulus covers the receptive field and therefore clockwise inward spiral will evoke the strongest response. In location 2 contraction would evoke the strongest response. In location 5 expansion or clockwise outward spiral would evoke strong responses. Thus, the cell might appear to be tuned for spiral motion (although not position invariant), but the tuning would be weak because the stimulus never fully matches the neuron's true preference and therefore evokes weaker responses.

stimulus location, and spatial response profiles with multiple peaks. Perge and colleagues (2005) used a motion reverse correlation approach (see Borghuis et al., 2003, for methodological details) to study the temporal dynamics of direction sensitivity in MT and found a temporal biphasic filter with neurons responding best when motion in the antipreferred direction was followed by motion in the preferred direction. Finally, Chen et al. (2008) used an approach not unlike ours to obtain receptive fields maps for MSTd neurons. Their stimulus consisted of a 4×4 or 6×6 grid of subfields, each of which contained dots moving in one of 8 directions that were updated every 100 ms. Thus, the stimulus was much lower in dimensionality compared to ours, which has 10×15 (or 6×9) segments and dots moving in any random direction. They used this approach mainly to compare direction tuning and spatial receptive field profiles between single unit and multiunit activity to study the clustering of tuning properties.

In contrast to these previous attempts, the purpose of our experiments was to explore whether reverse correlation can be used to characterize MST receptive fields in a way that allows to account for the tuning to complex motion patterns and for position invariance while possibly revealing additional features of the response profile that could not be detected with a more restricted stimulus set. This is not trivial, as position invariant responses to spiral motion patterns suggest that neurons respond to global motion patterns and do not integrate linearly across local patches (see Wild & Treue, 2021b, for a detailed discussion of the issue). The STA treats every segment of our RC stimulus as independent and does not consider dependencies between neighboring segments. It is therefore not surprising, that this approach is limited in its ability to predict responses to spiral stimuli across different locations (Fig. 5.6 E-H). Another problem is the large number of parameters that define our stimulus (up to 10×15 local direction seeds, resulting in a 150-dimensional vector of directions) and serve as input for the models. Regularization, which we included in our 4th model (Poisson GLM with regularization), is typically used to address such as "parameter explosion" by reducing weights in order to avoid overfitting and to reduce error rates in high-dimensional inference problems (Park & Pillow, 2011). In our study, one would expect this to bias weights outside the spatial receptive field towards zero. However, our tests of the models performance relied on predicting responses to stimuli that had been presented inside the receptive field, so that any reduction in weights outside the receptive field would have little effect on our performance measures. This could explain why there was no improvement in the performance of model 4 (light blue in Fig. 5.6E-H). The model that is most different from the others is model 5 (Spike-triggered covariance model), which incorporates STC, a measure that does take the covariance between

stimulus segments into consideration. We arbitrarily decided to extract 3 filters for each cell because the additional variance explained by each additional filter falls off quickly (Rust et al., 2005; Schwartz et al., 2006; Touryan et al., 2002). The fact that this did not lead to any critical improvement compared to the other four models (dark blue points in Fig. 5.6E-H) suggests that even more filters would not have made a meaningful difference, as they would have contributed even less.

An alternative approach are hierarchical models that describe a neurons response not as a function of the stimulus, but instead filter the stimulus with multiple subunits and then combine the output of these subunits to predict the neurons response. Such an approach has been successful in modeling responses of position invariant complex cells in primate V1 (Vintch et al., 2015) as well as responses of MST neurons (Mineault et al., 2012). The study by Mineault and colleagues (2012) found MST responses were much better predicted when the output of the subunits (which were modeled to be similar to MT neurons) was transformed by a nonlinear operation before integrating them. This evidence for the importance of nonlinear integration offers another possible explanation why our approaches, which are all inherently linear, were not more successful.

A potential reason why our approach was not more successful is that our stimulus is very complex and high-dimensional. As shown by the "Tuning" experiment, MST neurons respond very strongly to linear and spiral motion; the chance that our "RC stimulus" would randomly show a motion pattern that resembles any of these more structured stimuli is very low and even if it did happen, it would be hidden between thousands of other, unstructured samples. A more fruitful approach for future experiments could be to find a middle ground between the highly limited stimulus set used in the "Tuning" experiment and the high-dimensional, random stimulus used in the "Reverse Correlation" experiment. Adaptive, "closed-loop" methods, in which the neural responses are analyzed "online" as the experiment is going on and stimuli are modulated based on what has already been learned about the cell offer such an alternative (Benda, Gollisch, Machens, & Herz, 2007; Lewi, Butera, & Paninski, 2009) and have successfully been employed in the ventral visual pathway (Yamane et al., 2008).

Conclusion

We provide a detailed description of the response properties of neurons in macaque extrastriate area MST. Most of these neurons have large, but still clearly defined receptive fields and are strongly tuned to linear and spiral motion. A majority also showed position invariant tuning for spiral motion,

suggesting that this tuning is not a byproduct of improper stimulus placement. However, analyzing responses to a restricted stimulus set always runs the risk of missing important aspects of a neuron's response profile. Regression-based analyses of responses to high-dimensional, random stimuli are less biased and can detect previously unknown features of a response profile. Combining such methods with a novel random motion stimulus allowed us to predict some of the tuned responses of MST neurons, albeit generally less successful than in earlier visual areas. Thus, future experiments would benefit from the combination of unbiased high-dimensional stimuli and adaptive sampling methods that take into consideration what is already known about a cells tuning to avoid the "parameter explosion".

Chapter 6

Characterizing MST receptive fields with adaptive sampling

Benedict Wild & Stefan Treue

in preparation

As the previous chapter showed, regression methods that have been very successful in describing response properties of neurons in the early visual system were less effective when applied to MST neurons. This chapter provides an alternative approach for characterizing receptive fields of neurons in a high-level visual area such as MST.

Author contributions

B.W. and S.T. designed research. B.W. performed research. B.W. analyzed data. B.W. wrote the manuscript. S.T. supervised the project.

Acknowledgements

We thank Leonore Burchardt and Sina Plümer for assistance in animal handling, training, and surgeries, the German Primate Center's veterinary and animal husbandry staff for their expert animal care, Akshay Edathodathil for help with data collection, and Ralf Brockhausen for technical support, especially with regard to the hex-stimulus and communication between MWorks and Matlab.

Abstract

Characterizing the receptive fields of neurons in high-level visual areas poses a number of challenges. Neuronal activity in these areas typically depends on the interactions of many different stimulus features, which span a stimulus space that is too large to sample completely. Adaptive sampling methods offer a way of exploring this high-dimensional stimulus space in an efficient manner that makes use of prior knowledge about a neuron's tuning properties. We present two tools that can be used in such an adaptive sampling approach to characterizing the response properties of neurons in the motion sensitive medial superior temporal area (MST) of the primate visual cortex. The first tool is a newly designed random dot pattern that consists of hexgonal segments, each of which can be assigned independent values for direction of motion, speed, number of dots, and coherence. We compare neuronal responses to this stimulus to responses to conventional random dot patterns and find them to be very similar. The second tool is an implementation of the Nelder-Mead optimization algorithm for an online analysis of neurophysiological recordings in awake, behaving nonhuman primates. In simulations we show that this algorithm estimates a neuron's preferred stimulus in as few as 15 trials. The combination of these two tools offers great potential in the exploration of MST responses by providing information about neuronal characteristics in a more efficient, detailed and reliable way.

6.1 Introduction

The visual system of the primate cortex is hierarchically organized (Felleman & Van Essen, 1991). As information ascends through this hierarchy of different brain areas, the representation in each of these areas becomes less and less determined by physical attributes of the external world, and more and more an abstract, perceptual interpretation (Treue, 2003). This poses significant challenges for describing how responses of neurons in the later stages of this processing hierarchy depend on external stimuli. While neurons in early visual cortex (retinal ganglion cells, lateral geniculate nucleus, and primary visual cortex, V1) can be well-described as linear filters (e.g., Rodieck, 1965), possibly with the addition of a simple nonlinear function of the filter output (e.g., (Heeger, 1992)), neurons in high-level areas, such as the inferotemporal (IT) and the medial superior temporal (MST) cortex, elude such a simple description. Cells in IT cortex have large receptive fields, often extending into the ipsilateral hemifield, and respond strongly to complex shapes, such as objects, scenes, body parts, and faces (Gross, Rocha-Miranda, & Bender, 1972; Desimone, Albright, Gross, & Bruce, 1984; Hung, Kreiman, Poggio, & DiCarlo, 2005; Freiwald & Tsao, 2010; Bao et al., 2020, see Gross, 2008, and DiCarlo et al., 2012, for reviews). Cells in area MST have similarly large receptive fields and respond selectively to complex motion patterns, including the kind of optic flow motion that plays an important role in self-motion perception (Graziano et al., 1994; Duffy & Wurtz, 1995, see Wild & Treue, 2021b (chapter 2), for a review). Importantly, neurons in both areas are more selective for the identity of a stimulus, but also more "tolerant" or "invariant" for changes in low-level stimulus features than their respective input areas (MT for MST, V4 for IT) (Lagae et al., 1994; Rust & DiCarlo, 2010), which makes it difficult to describe their response as a function of such stimulus features.

A promising approach to address this issue are so-called *adaptive sampling* methods, in which data is analyzed as it is collected and stimuli for subsequent trials are chosen based on the responses to previous trials (Benda et al., 2007). Such an on-line analysis can be computationally quite expensive, however, and has therefore become a viable method only in recent years. An early suggestion for such an approach in the domain of visual neuroscience was the *Alopex* algorithm (Harth & Tzanakou, 1974; Tzanakou et al., 1979), which updates each pixel across the screen based on whether the neural response has changed consistently with the pixel's luminance in the previous two iterations. This suggestion was rarely adapted in actual experiments, however (see Micheli-Tzanakou, 1983, for an exception). In the auditory domain, Nelken and colleagues (1994) used the Nelder-Mead simplex method (Nelder & Mead,

1965; Press et al., 2002), which will be described in detail below, to find the combination of frequencies in a complex tone that are most efficient in driving neurons in cat primary auditory cortex. Yamane et al. (2008) developed an "evolutionary stimulus strategy" to investigate how complex 3D shapes are encoded by neurons in primate IT cortex. Starting out with an initial "generation" of 50 random 3D shapes, the probability that each member of this generation would produce probabilistically morphed descendants depended on the neuronal response to the "parent".

All these methods are optimization procedures that try to find a single "best" stimulus or a highly effective subspace of the virtually infinite stimulus space. In contrast, a method based on maximizing the mutual information between stimulus and response can estimate an optimal distribution of inputs (Machens, 2002). This approach was successfully used in the grasshopper auditory system to search for stimuli that could be described by a two-dimensional probability distribution over the mean and standard deviation of the amplitude of white noise snippets (Machens et al., 2005). In a similar vein, Lewi et al. (2009) used Generalized Linear Models (GLMs) to describe a neuron's response as a function of the current stimulus, the stimulus history, and the neuron's spiking history. They also maximize the mutual information between the collected data and the parameters of the GLM and rely on approximations to make this process computationally efficient and feasible for real experiments.

However, such methods become more difficult to apply if the neuronal response depends on multiple parameters with different distributions and constraints and different functional relations to the neuronal response. MST neurons, for example, are tuned for location (through their spatial receptive field), motion direction, speed, binocular disparity, and possibly other, as of yet unknown variables. Their tuning for location and direction can be described by a 2D and circular Gaussian respectively, which are well suited for many estimation methods. The tuning for speed, however, cannot easily be described by such a "well behaved" function and has been approximated using smoothing splines (e.g., A. K. Churchland et al., 2007; also see chapter 3). It is likely because of this complexity that no one has so far actually implemented such model-based approaches to characterize the response properties of high-level visual areas in awake, behaving nonhuman primates.

Therefore, we focus here, as a first step, on the simpler optimization approach and present tools that allow to quickly find the optimal stimulus that maximally drives a MST neuron. The first of these tools is a newly developed stimulus that is made up of individual, hexagonal segments, which allows for control of local aspects, but can also create coherent motion patterns, similar to those that are typically shown with classic random dot patterns (see

chapter3). The second tool is an implementation of the Nelder-Mead simplex method for a receptive field mapping experiment in awake, behaving animals. Due to technical issues we have not yet been able to test the combination of these two tools in animals. We therefore separately present preliminary results from recordings with the new stimulus, as well as computer simulations of the adaptive sampling approach.

6.2 Methods

6.2.1 Physiological recordings

Subject

One male rhesus monkey (*Macaca mulatta*) was tested in the experiments. The monkey was 10-11 years old and weighed between 9.8 and 11.4 kg during the period of data collection. He was implanted with a titanium head holder to minimize head movements during the experiment, as well as with a recording chamber implanted on top of a craniotomy over the left parietal lobe based on a magnetic resonance imaging (MRI) scan. Surgeries were conducted under general anesthesia and post-surgical care using standard techniques.

Task and neural data acquisition

Both the task and the neural recording procedures were identical to those described in chapters 4 and 5. Briefly, the monkey was required to keep its gaze within a small window around a fixation dots and respond to a change in the fixation dot's luminance by releasing a button. Spikes were recorded single tungsten electrodes or tetrodes (Thomas Recording, Giessen, Germany) that were advanced into the recording chamber using a mechanical microdrive ("Mini Matrix", Thomas Recording, Giessen, Germany). The recorded signal was digitized, filtered, and spike waveforms were manually sorted into clusters.

Experiments

We ran the same "Spatial Mapping" and "Tuning" experiments described in chapter 3 as a "ground truth" of each recorded cell's response properties. The only differences to the procedures described previously were that in the "Tuning" experiment, spiral motion, like linear motion, was presented in only one location and the probe duration lasted 167 ms instead of 100 ms.

In addition, we recorded responses to a newly developed random dot pattern that consists of hexagonal segments ("hex-stimulus", Fig. 6.1). The

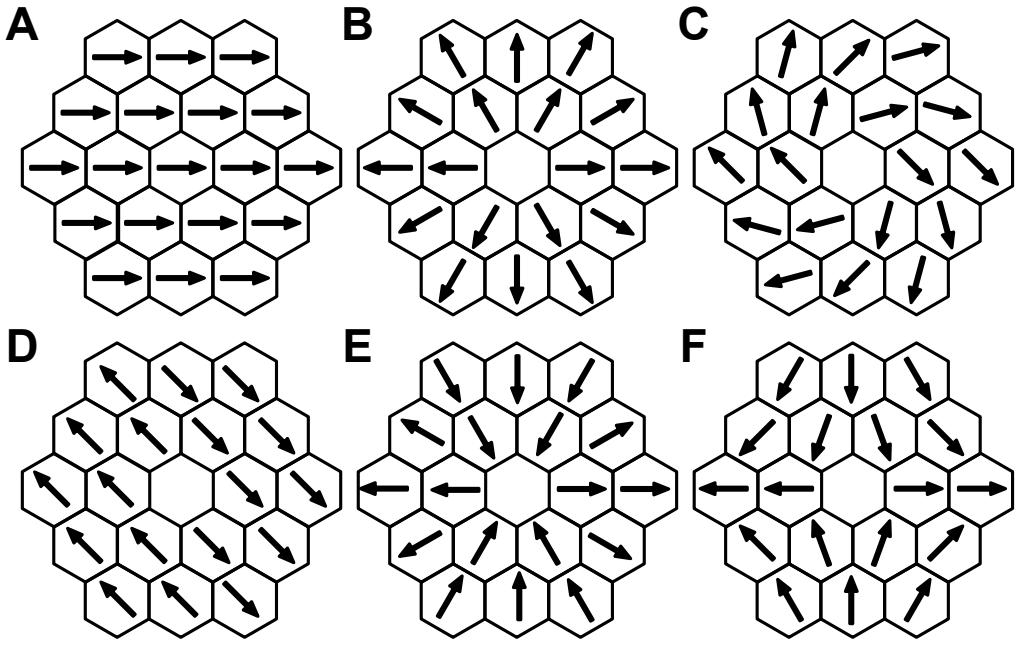

Figure 6.1: Six example motion patterns of the newly developed hex-stimulus. ***A***. linear rightward; ***B***. expansion; ***C***. outward clockwise spiral; ***D***. shearing; ***E***. deformation; ***F***. saddle point.

motivation behind this stimulus was two-fold: on the one hand, a stimulus consisting of individual segments offers fine-grained control of local features; on the other hand, using hexagons allows to generate the kind of motion patterns that are typically used to probe MST neurons, such as radial and rotational motion. The hex-stimulus combines both of these features: by using more and more, smaller and smaller hexagons, one can create arbitrarily complex and random motion patterns, but the motion in every segment is under the control of the experimenter. At the same time, already 18 hexagons, arranged in two concentric circles, are sufficient to create motion patterns that resemble expansion, contraction, rotation, or spiral motion very well and can also create additional patterns, such as shearing and deformation (Fig. 6.1.

For the data presented here we used a version of the stimulus with 19 hexagonal segments arranged in 2 concentric circles and created a total of 22 motion patterns:

- 8 linear motion patterns in which all 19 segments showed dots moving in the same direction (0° to 315° in steps of 45°) with the same speed

(e.g., Fig. 6.1A).

- 8 spiral motion patterns: expansion (Fig. 6.1B), outward clockwise spiral (Fig. 6.1C), clockwise rotation, inward clockwise spiral, contraction, inward counterclockwise spiral, counterclockwise rotation, and outward counterclockwise.
- 2 shearing motion patterns (Fig. 6.1D as well as the opposite directions).
- 2 deformation motion patterns, in which dots move towards the center along either the horizontal or the vertical axis and away from the center along the other axis (Fig. 6.1E).
- 2 saddle point motion fields (e.g., Koenderink & van Doorn, 1975), in which the dots move towards the center along either the vertical (Fig. 6.1F) or the horizontal axis and then are "deflected" away from the center along the other axis.

Note that for all but the 8 linear motion patterns, the segment in the center of the stimulus was left blank (dot density of 0 dots/deg^2) because no local linear motion would fit the overall motion pattern. The stimulus was updated to a new configuration of location directions every 167 or 200 ms throughout the duration of a trial. In the early recording sessions only one or two speeds were tested, but in later recording sessions, the same speeds as in the "Tuning" experiment were used. For spiral motion, the 12 outermost segments moved with a faster speed to simulate the speed gradient that is used in spiral RDPs (see chapter 4).

6.2.2 Simulation

We simulate single MST neurons based on our measurements of real neurons. A simulated neuron is defined by

- A baseline firing rate (randomly chosen between 2 and 14 spikes/s).
- A maximum amplitude (randomly chosen between 10 and 100 spikes/s).
- A spatial receptive field, defined as a difference-of-two-2D-Gaussian with randomly chosen center coordinates, standard deviations (between 11° and 20° along one axis and between 9° and 18° along the other axis), and rotation angle. We used a difference-of-Gaussians surface to simulate an inhibitory surround (even though this is not necessarily the case for MST neurons, see Chapter 2), because otherwise the optimal

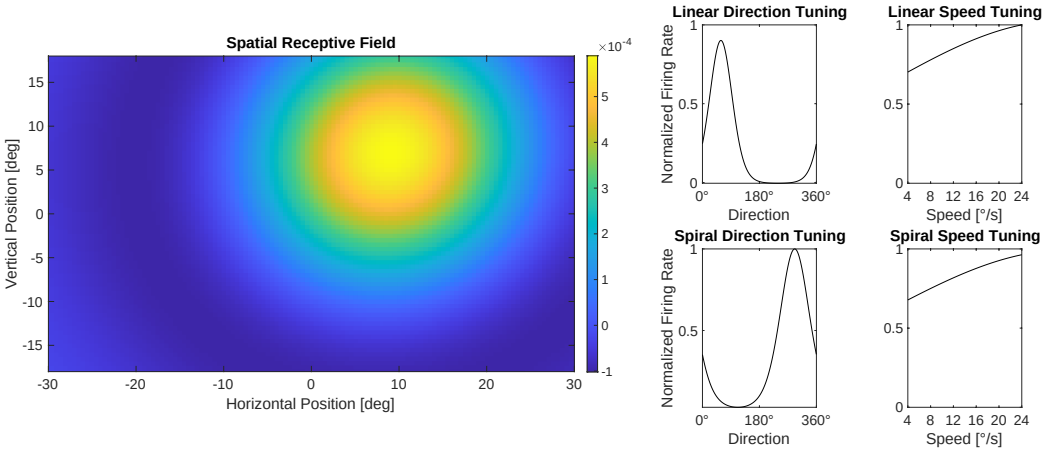

Figure 6.2: Example of the tuning functions that define our model neurons: each neuron has a spatial receptive field and direction and speed tuning curves in linear and spiral space. What is not shown in this illustration are the remaining two parameters that define a simulated neuron, it's baseline firing-rate and it's maximum amplitude.

stimulus size would have no upper bound. The "inhibitory surround" had the same coordinates and rotation angle as the "excitatory center", but standard deviations that were 20% larger and an amplitude that was 20% of the "excitatory center"

- Two circular Gaussian direction tuning curves (one in linear and one in spiral motion space);
- Two speed tuning curves (one in linear and one in spiral motion space).

The speed tuning curves are modeled as normal distributions with a mean chosen from a uniform distribution between 16 and 32 °/s. As we only define speed tuning in a space between 4 and 24 °/s, most simulated neurons show a monotonically increasing speed tuning (because the mean of the underlying normal distribution lies beyond the range of tested values) and only a small minority show a preferred speed slower than 24 °/s and decrease their response if the speeds exceeds this preferred speed, which is in accordance with our observations of real MST neurons (see Chapter 5). The spatial response functions is normalized so that all values sum to 1 and the direction and speed response functions are normalized to values between 0 and 1. An example for the spatial receptive field and direction and speed tuning curves that define our model neurons is shown in Figure 6.2.

A stimulus is defined by an x- and y-coordinate, a radius, a direction, a

speed, its motion space (linear, spiral, or non-coherent), and its duration (in seconds).

The response r of a neuron to stimulus presentation is drawn from a Poisson distribution with rate parameter λ.

$$r = Poiss(\lambda)$$

This parameter, which can be interpreted as the neuron's average response to that stimulus, is defined as the neuron's baseline firing rate plus a product of the maximum amplitude A and a scaling factor α:

$$\lambda = FR_{baseline} + \alpha A \qquad \text{with } \alpha \in [0, 1]$$

This scaling factor consists of a spatial ($\alpha_{spatial}$), a directional ($\alpha_{directional}$), and a speed (α_{speed}) component.

The spatial component is the sum of the part of the spatial receptive field that is covered by the stimulus. Because this value was so low for small stimuli that it led to unrealistically low firing rates, the square root is taken (which, as it is a fraction between 0 and 1, scales it up).

$$\alpha_{spatial} = \sqrt{\iint f(x, y) I(x, y)\, dx\, dy}$$

where f is the receptive field and I the stimulus

The directional and speed compoenents are simply the values of the direction and speed tuning curve (which are already normalized to range from 0 to 1). The three components are multiplied and the square root is taken again to avoid unrealistically low firing rates.

$$\alpha = \sqrt{\alpha_{sptial} * \alpha_{directional} * \alpha_{speed}}$$

To this rate parameter, random noise with mean 0 and standard deviation $\frac{\lambda}{2}$ is added[1]. Finally, the resulting value is half-wave rectified to avoid a negative response.

All these values and procedures were chosen so as to result in simulated data that was similar to the recorded data.

Simulated mapping and tuning experiments

To test whether our modeled neuron was able to generate responses similar to real neurons, we simulated both the "Spatial Mapping" and the "Tuning"

[1]We found that this made the simulated data more similar to real data in terms of noisiness

experiment. As we wanted to compare whether our adaptive approach can lead to similar results in fewer trials, we simulated all experiments with "trials" with a random length between 3.5 and 4.5 s in which samples were sequentially presented, as in the real experiments. After every sample there was a 5% chance of a "fixation break" that immediately ended the trial to account for the fact that in real recordings, trials that end in fixation breaks by the monkey are shorter and provide fewer samples. As in the physiological recordings, a maximum number of samples (6000 for "Spatial Mapping" and 1200 for "Tuning") was set and "trials" were simulated until this number of samples was reached.

Simplex optimization

The Nelder-Mead simplex algorithm was originally developed to minimize functions with n independent variables without the need for derivatives (Nelder & Mead, 1965; Press et al., 2002). This makes it well suited for our purpose where we try to maximize[2] an MST neuron's firing rate which we assume to be a function of stimulus parameters, such as location, direction, and speed. The basic algorithm as originally described works like this (translated to maximization):

1. To maximize a function with n parameters, choose $(n + 1)$ points P_0, P_1, ..., P_n in n-dimensional space that form a simplex (i.e., a triangle in 2D, a tetrahedron in 3D, etc.).

2. P^* is the *reflection* of the point that has the lowest function value (P_l) across the centroid of all other points. If the function value of P^* lies between the function values of P_l and the point with the highest function value, P_h, then replace P_l with P^*.

3. If the function value of P^* is higher than the function value of P_h, then *expand* P^* to P^{**} (i.e., P_l is reflected across the centroid and expanded by the same distance again). If P^{**} is larger than P_h, replace P_l with P^{**}. Otherwise replace P_l with P^*.

4. If the function value of P^* is smaller than all other P except P_l (i.e., P^* would be the worst point of the new simplex), *contract* P_l to P' (i.e., the point halfway between P_l and the centroid of all other points). If the function value of P' is larger than that of P_l, replace P_l with P'.

[2]Whereas Nelder and Mead originally described the algorithm to minimize a function, we want to maximize a function, which trivially means exchanging "highest" for "lowest" and vice versa in all steps of the algorithm

5. If the P^*, P^{**}, and P' are all smaller than the smallest point of the simplex, shrink the simplex by keeping its highest point P_h and reducing the distance between all other points and P_h by 50%.

There are two difficulties in applying this algorithm to neural recordings from awake behaving animals. First, because the function we are trying to maximize (i.e., the neuron's response function) is subject to noise, we cannot simply "evaluate" it, but need to take the average of multiple "evaluations" (i.e., stimulus presentations). Second, the algorithm often requires conditional function evaluations (i.e., if $P^* > P_h$, then evaluate P^{**}). This is not trivial in a case where every "evaluation" (i.e., stimulus presentation) is expensive, in the sense that the overall number of "evaluations" we can perform depends on the animal's motivation to work and is, in fact, the very thing that we want to reduce.

We address these issues as follows: For every trial (which lasts 3500 to 4000 ms or until a fixation break) we follow the procedure from the "Spatial Mapping" and "Tuning" experiments and present a rapid succession of 100 ms samples that are randomly drawn from the $n + 1$ possible stimuli that define the current simplex. We keep drawing from the same stimuli until we have at least 10 samples for each one of the $n + 1$ stimuli, repeating the process across multiple trials, if necessary. Then we interrupt the normal sampling procedure to do an "REC"-trial (Reflection, Expansion, Contraction), in which we "evaluate" the function at all three potential points. We can then choose the appropriate stimulus from the REC-trial to replace the weakest stimulus of the previous regular trial and return to the regular sampling procedure. Figure 6.3 demonstrates the first three regular and first three REC trials of this procedure for the simple and graphically intuitive case of maximizing a visual neuron's firing rate as a function of 2 parameters: the x- and y-coordinate of a stimulus (similar to our "Spatial Mapping" experiment). Because there are 2 parameters we use 3 stimuli (i.e., 3 combinations of a x- and a y-coordinate) that are initially placed randomly (Trial 1 in Fig. 6.3). The stimulus in each regular trial that elicited the weakest response (marked in blue) needs to be replaced. On REC trials a reflected, an expanded, and a contracted option are presented and the option that will replace the weakest stimulus from the previous regular trial is selected based on the rules of the algorithm outlined above (selected option is marked in cyan in Fig. 6.3).

6.3 Results

6.3.1 Physiological recordings

We analyzed 54 cells that had been recorded in 50 sessions. Out of these 54 cells, one was excluded from all analyses because it did not provide inter-

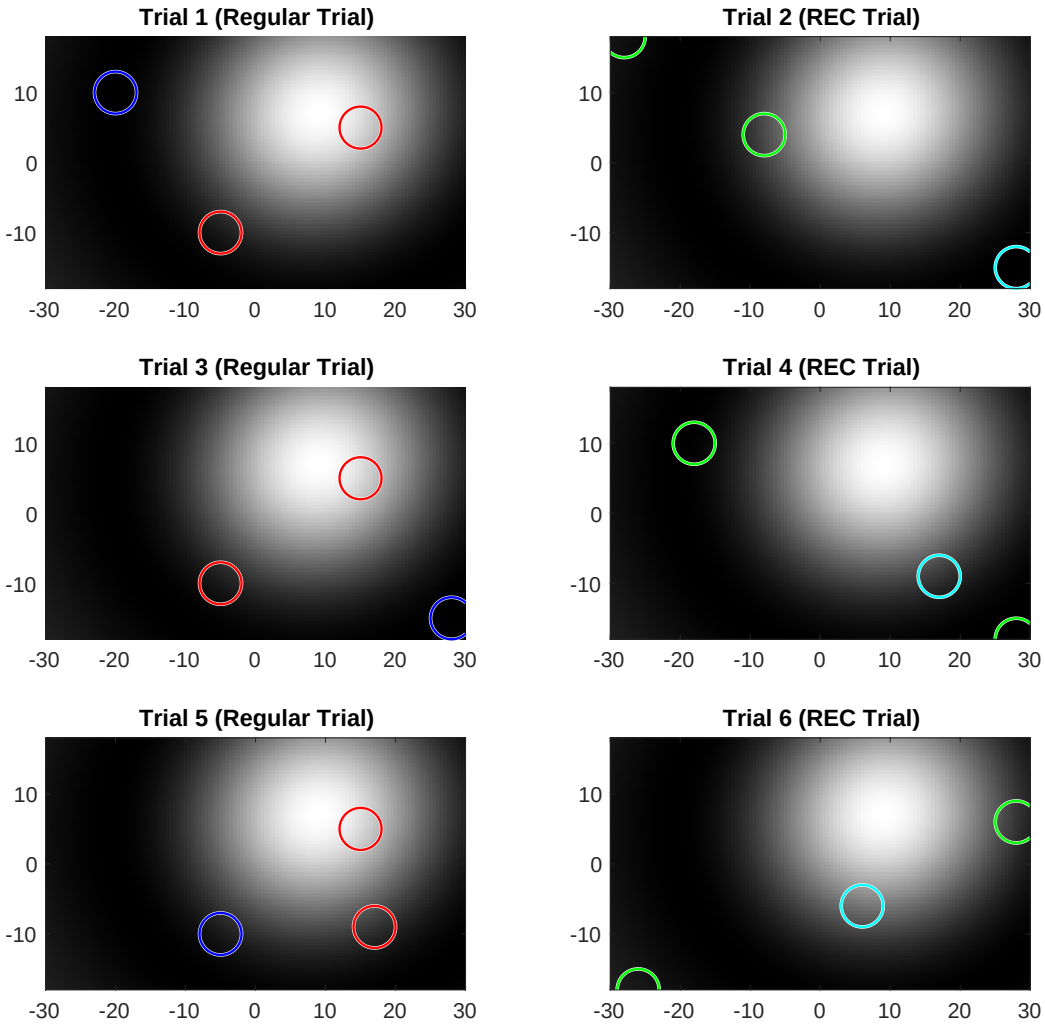

Figure 6.3: The first 3 regular (left column) and REC (right column) trials for a simulation of the simplex method with 2 parameters (x- and y-coordinate). The simulated neuron's spatial receptive field is shown as a grayscale map in the background. In the regular trials, the blue stimulus is the one that elicited the weakest response and needs to be replaced. In the REC trials, the cyan stimulus is the one that is selected to replace the weakest stimulus from the previous regular trial.

pretable data in either the "Spatial Mapping", nor the "Tuning experiment" (i.e., it did not have a clear spatial receptive field and showed no tuning for linear or spiral direction).

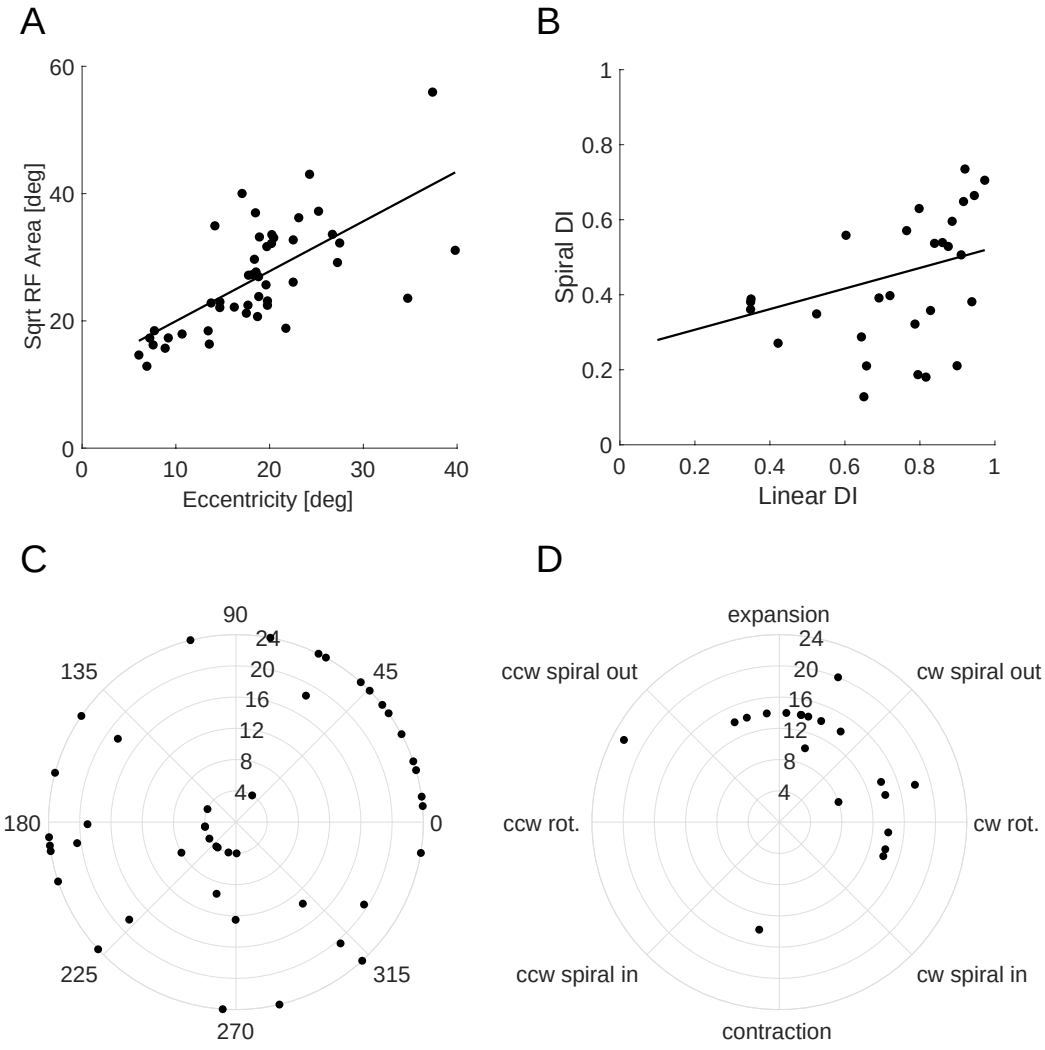

Figure 6.4: Population results for "Spatial Mapping" and "Tuning" experiment. ***A***. Receptive field size (square-root of the area) as a function of eccentricity across all cells. ***B***. Scatter plot of linear and spiral directionality index. ***C***. Distribution of preferred directions (angle) and speeds (radius) across all cells for linear motion. ***D***. Like C, but for spiral motion.

Spatial Mapping and Tuning

We followed the same analysis steps as described in chapter 5. Out of 53 cells that entered the analysis, 3 were excluded from the "Spatial Mapping" analysis because the coefficient of determination (R^2) of the fit was less than 0.22, and another 4 cells were excluded because their receptive field center lay more than 2 degrees outside the screen. For the remaining 46 cells, we found a highly significant correlation between eccentricity and receptive field size (quantified by the square-root of the area) ($r = .68, p < .001$), similar to the data described in chapter 5. The slope of a linear regression line was 0.78 with an offset of 12.1 (Fig. 6.4).

For the "Tuning" experiment, 7 cells were excluded because they were tuned neither for linear, nor for spiral motion (see chapter 5 for criteria). Of the remaining 46 cells, 16 were tuned only for linear motion, 2 were tuned only for spiral motion, and 28 were tuned for both, linear and spiral motion. Figure 6.4C and D show the distribution of preferred directions and speeds for linear (panel C) and spiral (panel D) motion. As for the data described in chapter 3, the distribution of preferred linear directions did not differ significantly from a uniform distribution ($p = .78$, Hodges-Ajne test for non-uniformity of circular data), but the distribution of preferred spiral directions did differ significantly from a uniform distribution ($p < .001$) with a bias towards outward spiral motion (average preferred direction: 41.25° where 0° is clockwise rotation and 90° is expansion, Fig. 6.4D). Figure 6.4B shows the relation of linear and spiral tuning strength, quantified through a directionality index ($DI = 1 - \frac{R_{pref}}{R_{anti-pref}}$). Again, tuning strength for linear and spiral motion was significantly correlated ($r = 0.43, p < .05$).

Hex-stimulus

Our analysis of the neuronal responses to the newly developed hex-stimulus had two goals: First, we wanted to test whether motion patterns created with the hex-stimulus that resemble linear and spiral motion (e.g., Fig. 6.1A-C) evoke similar responses as classic random dot patterns (RDPs). Second, we wanted to test whether neurons' responses to other motion patterns, such as shearing and deformation (e.g., Fig. 6.1D-F) are of a similar magnitude or even in excess of responses to linear and spiral motion. This would suggest that the underlying response function of these neurons is more complex than two tuning curves in linear and spiral motion space.

Figure 6.5 shows responses of one example neuron to the RDPs used in the "Tuning" experiment (panels A and B), responses to motion patterns generated

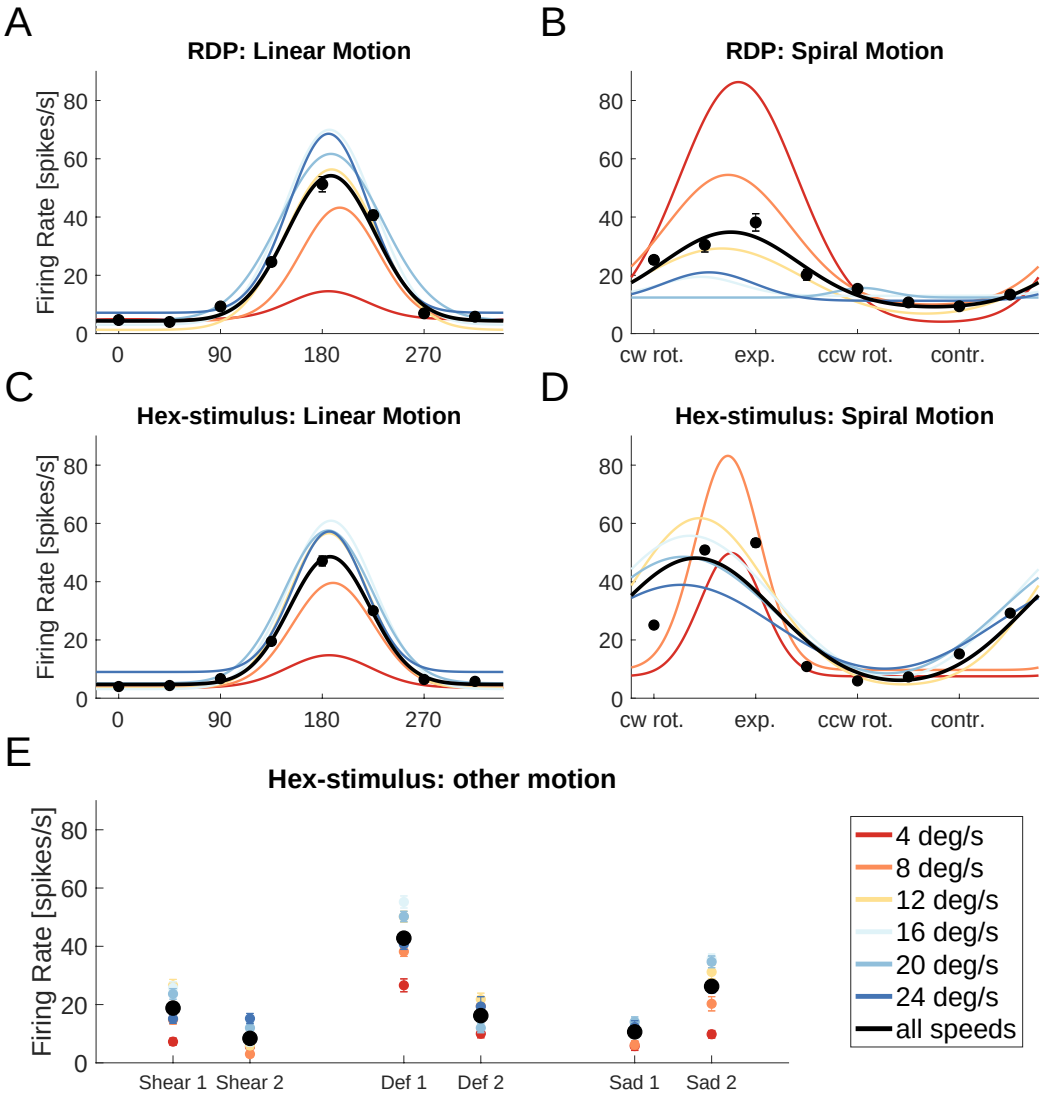

Figure 6.5: Comparison of an example neuron's responses to the RDPs from the "Tuning" experiment and the hex stimulus. ***A***. Direction tuning curve for linear motion with RDPs. ***B***. Like A, but for spiral motion. ***C***. Direction tuning curve for linear motion with hex-stimulus. ***D***. Like C, but for spiral motion. ***E***. Responses to other motion patterns tested with the hex stimulus: Shear 1 (Fig. 6.1D) and Shear 2 (opposite motion of Fig. 6.1D); Deformation (Def) 1 (Fig. 6.1E) and Deformation 2 (opposite motion of Fig. 6.1E); and Saddle point motion (Sad) 1 (Fig. 6.1F) and Saddle point motion 2 (opposite motion of Fig. 6.1F)

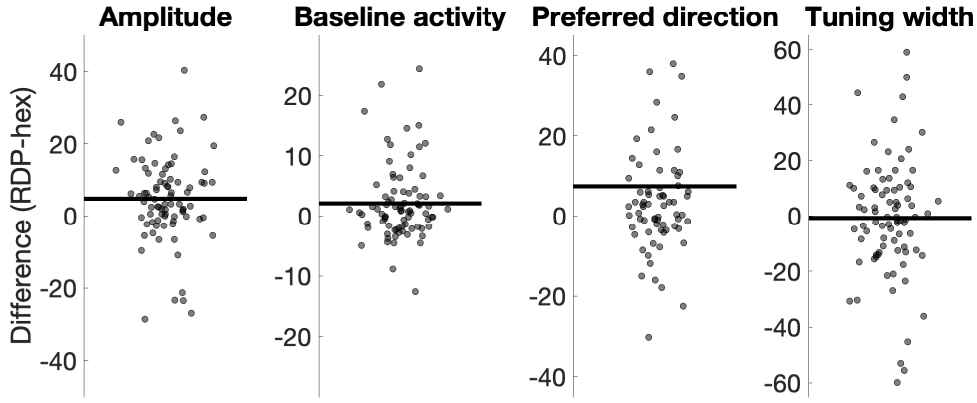

Figure 6.6: Comparison of tuning curve parameters for curves fitted to responses to RDPs and the hex-stimulus. The 83 points in each plot show the individual comparisons, the thick lines show the mean of the distribution.

with the hex-stimulus that try to simulate these these RDPs (panels C and D), and responses to 6 additional motion patterns that were not part of the "Tuning" experiment (two types of shearing motion, two types of deformation motion, and two types of saddle point motion, panel E). For linear motion (panels A and C), the responses are almost indistinguishable. For spiral motion (panels B and D), response patterns differ a bit more, but the tuning with a preference for expansion is present for both stimulus types. Finally, the neuron responds to deformation motion type 1 with responses that are similar in magnitude to those for the preferred linear direction. Thus, the data of this example neuron suggest that motions patterns created with the hex-stimulus do evoke similar responses as RDPs, and that motion patterns outside the linear and spiral motion space can also evoke strong responses, hinting at a more complex underlying response function.

To compare responses to RDPs and the hex-stimulus across the population, we compared the four parameters that define each tuning curve (baseline firing rate, amplitude, preferred direction, and tuning width) for each pair of RDP and hex-stimulus tuning curves in which both curves met our inclusion criteria across 15 cells for which we had tested both stimulus types with the same speed values. Out of a total of 180 pairs of tuning curves (15 cells with tuning curves for 6 speeds for both linear and spiral motion), 83 pairs of tuning curves (46%) entered into the comparison. For 11 pairs, only the responses to the RDP resulted in satisfactory tuning curves, for 28 pairs, only the responses to the hex-stimulus resulted in satisfactory tuning curves, and for 58 pairs responses to neither stimulus resulted in satisfactory tuning curves. Thus, surprisingly,

responses to the hex-stimulus could be better described with tuning curves than the responses to the RDP. The comparison of the four tuning curve parameters for 83 curves is shown in Figure 6.6. We always subtracted values from curves fitted to hex-stimulus data from values from curves fitted to RDP data, so that values larger than 0 mean a higher baseline firing rate, or larger amplitude, or wider tuning width, or a preferred direction shifted in anticlockwise direction for responses to the RDP as compared to the hex-stimulus. We performed 4 t-tests with a Bonferroni-corrected significance level α of $0.05/4 = 0.0125$ to test whether each distribution differed significantly from 0. Differences in preferred direction ($t(82) = 0.67, p = .51$) and tuning width ($t(82) = -0.33, p = .74$) did not differ from 0. Differences in amplitude ($t(82) = 3.71, p < .01$) and baseline firing rate ($t(82) = 3.1, p < .01$) did differ significantly from 0. The result for the baseline firing rate is surprising because it was calculated based on spiking activity in the 200 ms before stimulus onset on each trial and there should not differ between the two stimuli. This suggests that the difference in baseline activity (in possibly in amplitude as well) is not caused by the different stimuli. Instead, it could be explained by a general decline in firing rate or in signal quality throughout the recording session, because responses to the hex-stimulus were always recorded after the "Tuning" experiment.

All together these results provide strong evidence that motion patterns created with the hex-stimulus evoke similar responses as random dot patterns.

To compare responses to the six new motion patterns with responses to linear and spiral motion across the population, we calculated the average response with a 95% confidence interval (1.96× standard error of the mean) and define a significant difference in responses to two motion patterns as non-overlapping confidence intervals. We had responses to all 22 motion patterns from 45 cells (30 with 1 speed and 15 with 6 speeds) and only compared motion patterns with the same speed, for a total of 120 comparisons (15 cells with 6 speeds plus 30 cells with one speed). In 2 of these cells, at least one of the six new patterns elicited a significantly larger response than the largest response to one of the 16 linear and spiral patterns. In 16 of these cells, at least one of the new patterns elicited a response larger than the 4th largest response to linear and spiral motion. And in 33 cells, at least one of the new patterns elicited a response larger than the median (i.e., the 8th largest) of the 16 responses to linear and spiral patters.

While further analyses are necessary to investigate these response behaviors in more detail (see also Discussion), this strongly suggests that MST neurons not only respond selectively to linear and spiral motion, but also to more complex patterns, such as shearing, deformation, and saddle point motion.

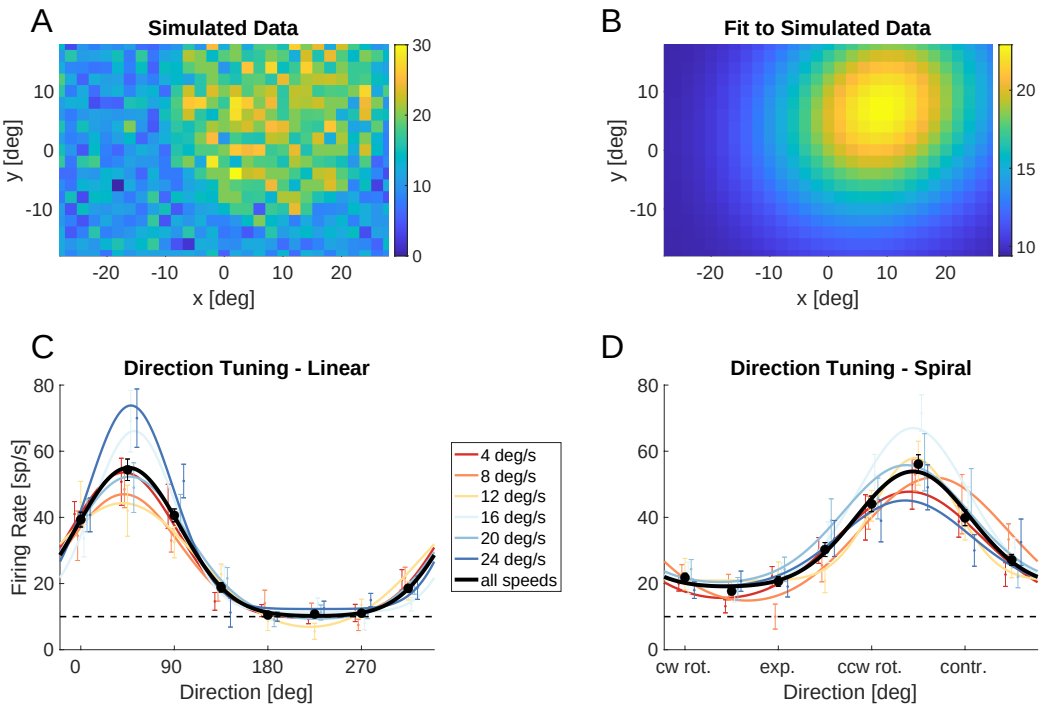

Figure 6.7: Results for the simulated "Spatial Mapping" and "Tuning" experiments with the model neuron that was depicted in Fig. 6.2. ***A***. Receptive field map. ***B***. 2D Gaussian fit to the data in panel A. ***C***. Tuning curves for linear motion. Color indicated different speeds. ***D***. Like panel C, but for spiral motion.

6.3.2 Simulation

To test whether the model is able to replicate the data recorded from real MST neurons, we repeated the "Spatial Mapping" and "Tuning" experiments with a simulated neuron. Fig. 6.7 shows the results. Based on visual inspection, the noise level of the spatial receptive field map (Fig. 6.7A) is similar to that of real neurons, suggesting that the parameters we selected result in responses that are comparable to real neurons. The 2D Gaussian fit (Fig. 6.7B) resulted in estimates of the receptive field parameters that were all very close to the real parameters (Table 6.1).

For the tuning experiment, the fitted tuning curves (Fig. 6.7C and D) described the preferred direction very well, but overestimated the tuning width (Table 6.1). Overall these results indicate that our model neuron's responses are comparable to real neurons and it is therefore well suited to explore the optimization algorithm.

Parameter	**Real value**	**Estimated value**
Spatial Mapping		
X-coordinate	9	9.015
Y-coordinate	7	7.039
Std. Dev. 1	14	13.57
Std. Dev. 2	13	11.68
Tuning		
Preferred linear direction	45	46
Linear tuning width	35	49
Preferred spiral direction	225	220
Spiral tuning width	45	55

Table 6.1: Real parameters of the simulated neuron and estimated parameters based on the simulated "Spatial Mapping" and "Tuning" experiments.

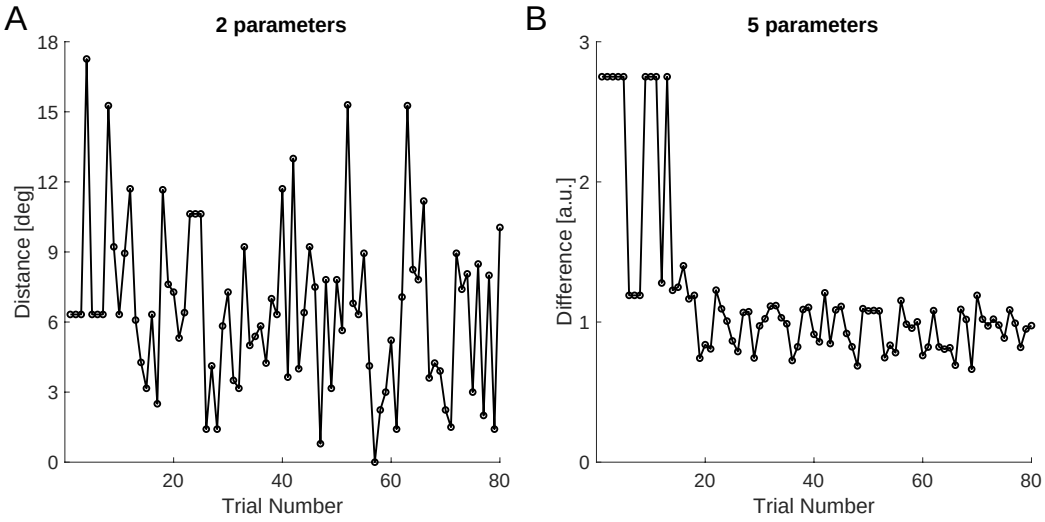

Figure 6.8: Difference between the stimulus that elicits the strongest response and the true parameters as a function of trial number for the simplex optimization simulation. ***A***. Simulation with 2 parameters (x- and y-coordinate); y-axis shows distance from best stimulus to true receptive field center in degrees. ***B***. Simulation with 5 parameters (x-coordinate, y-coordinate, radius, direction, speed); y-axis shows summed normalized difference between best stimulus and true values, which can range from 0 to 5.

We first ran the implementation of the Nelder-Mead simplex optimization on the 2D example described in Methods, where we attempt to maximize the model neuron's firing rate as a function of stimulus position. The radius

(2°), direction (incoherent), speed (randomly chosen for each sample), and sample duration (100 ms) of the stimulus remained the same and only x- and y-location were updated. This is most similar to the "Spatial Mapping" experiment. While the simulated "Spatial Mapping" experiment ran for 93 trials (which is comparable to the amount of trials used in the electrophysiology experiment), the simplex version of the experiment converged towards the center of the receptive field within as few as 10 trials (regular and REC trials combined). To quantify how close the algorithm came to the true receptive field center, we simulated 80 trials and after every trial (starting from trial 3) calculated the distance between the stimulus that had elicited the strongest average response up to that point and the receptive field center. As can be seen in Fig. 6.8A, the distance decreased to less than 3° within less than 20 trials but kept oscillating and reached values as large as 15° even after 60 trials. It should be noted that this model neuron's receptive field had standard deviations of 13° and 14° so that all distance values below 13° fall within one standard deviation of the 2D Gaussian receptive field, which can be considered a good estimate.

While this result is promising, it is of limited use. This approach allows a good estimate of the receptive field center (within one standard deviation of the true value) with few trials, but it offers no information about the spatial receptive field other than its center.

For the second simulation we attempted to optimize all 5 parameters that define our stimulus: x-coordinate, y-coordinate, radius, direction, and speed. Because our model, based on our findings in the physiological recordings, is tuned in both linear and spiral space with no connection between the two, we ran the simulation separately for linear and spiral motion. We normalized all values to range from 0 to 1 to ensure that changes in each parameter would be of similar size. Again, we ran the simulation for 80 trials and calculated the the sum of the normalized differences between the 5 attributes of the most effective stimulus and the true parameters of the model neuron after every trial (Fig. 6.8B). In contrast to the 2-parameter example, a clear drop in the difference can be observed after 15 trials. From then on the value oscillates around a stable, low value. This indicates that with as few as 15 trials, the algorithm comes very close to the optimal stimulus that is defined by 5 parameters. Note that with 5 parameters, several trials are necessary in the beginning of the simulated experiment until 10 samples have been collected for every stimulus configuration. After this initial data collection, however, the algorithm will switch between regular and REC trials on almost every trial, because only one stimulus is replaced after every iteration.

6.4 Discussion

Responses to the hex-stimulus

We have established a new motion stimulus, the *hex-stimulus*, that evokes responses similar to widely used random dot patterns, but also allows to create complex compositions of local motion. For the results we presented here, we only used a configuration of 19 hexagons arranged in 2 concentric circles with one hexagon in the center. This allowed us to simulate commonly used optic flow patterns, such as expansion and contraction, to a degree that they evoked similar responses as random dot patterns. While the overall motion pattern can be made even more similar to random dot patterns by using more and smaller hexagons, this does not seem to be necessary.

This stimulus opens a range of possibilities, only some of which have been explored here. First, it can be used to create complex motion patterns, such as shearing and deformation. We have tested only 6 such patterns, but by rotating the stimulus or by using more segments, even more complex patterns can be created. For the six patterns that we tested, we found that neurons responded strongly to them, suggesting that one dimensional tuning curves for translational or spiral motion might miss some of these neurons' truly preferred stimuli. Because we did not test for position invariance in this experiment, we cannot rule out that responses to the new motion patterns (shearing, deformation, saddle point motion) might have been an artifact of badly positioned stimuli. If, for example, the shearing pattern shown in Figure 6.1D is positioned so that only its upper right half falls into the receptive field, a neuron that is tuned for downward/rightward linear motion would respond very strongly. Future recordings should explore the responses to these new motion patterns in more detail to confirm whether responses are selective for these specific arrangements. Two previous papers also recorded MST neurons' responses to deformation Lagae et al., 1994; Mineault et al., 2012, but provided little information about how common selective responses to such stimuli are in a population of MST neurons.

A second possibility opened up by the hex-stimulus is to create "patchy" optic flow patterns by leaving some of the segments empty. This approach was used in a recent study that explored the ability of different primate species (human, macaque, and marmoset) to track the focus of expansion (FOE) in a large random dot pattern that also consisted of hexagonal segments (Knöll et al., 2018). Given that MST neurons appear to respond mostly to the overall, global motion pattern of a stimulus, rather than summing over local patterns, it would be interesting to explore to what degree a motion pattern can be degraded before a neuron stops showing tuned responses.

Finally, the stimulus can also be used to easily test different speed gradients, as the speed of each segment can be individually adjusted. A potential extension of the stimulus would be to add different binocular disparity levels, which MST neurons are also selective for (Roy et al., 1992).

Adaptive sampling

We presented an application of the Nelder-Mead simplex algorithm for characterizing the tuning of simulated MST neurons for up to 5 parameters. We were able to find a good approximation of a model neuron's preferred stimulus within approximately 15 trials. This is a huge saving in trials compared to the "Spatial Mapping" and "Tuning" experiments, which together take way beyond 100 trials to provide good data. However, as mentioned in the introduction, like all optimization algorithms, this one also only searches for the one single best stimulus. Often it is desirable, however, to know about the full response function. In the case of a direction tuning curve, for example, which is defined by at least two parameters (preferred direction and tuning width), the ability of a neuron to discriminate between two directions depends on its steepness, which makes the tuning width as important as the preferred direction. An optimization algorithm that attempts to maximize the neuron's firing rate will only search for the preferred direction. Furthermore this is problematic if there is not one single best stimulus. As a matter of fact this is the case for our simulated neuron, which has independent direction and speed tuning curves for linear and spiral motion. Nevertheless, the approach presented here offers valuable benefits. First, our simulation shows that a flexible approach in which trials are repeated until enough data has been collected to reduce noise (in our case: 10 repetitions per stimulus) is feasible. Second, with taking less than 20 trials to find a highly effective stimulus (even a single "best" stimulus might not exist), this approach can serve to quickly find a starting point for more elaborate procedures. Rhesus macaques can do several hundred trials per recording session for experiments like the ones described here and in chapters 4 and 5. Investing 20 of these trials to get a rough idea about a neuron's tuning can certainly be considered time well spent.

The interesting question is, of course, whether the optimization approach can be combined with the hex-stimulus. We did not test this here, because it would require a more complex model of neuronal responses than the one we used to generate responses to the hex-stimulus. However, we showed that 5 parameters can be optimized with as few as 15 trials. It seems reasonable to assume that a small version of the hex-stimulus (e.g., with only 7 segments arranged in a circle) where only the direction of each segment is fed into the

optimization algorithm should also converge quickly.

In conclusion, we present two tools – a new stimulus and an adaptive sampling approach – that can help to efficiently search the virtually infinite stimulus space that MST neurons respond to. These tools will be helpful to characterize the response properties of MST neurons more efficiently, and thus make it possible to investigate whether these properties are stable or depend on factors such as attention or adaptation.

Chapter 7

Comparing the influence of stimulus size and contrast on the perception of moving gratings and random dot patterns

Benedict Wild & Stefan Treue

in preparation

The previous chapters highlighted the importance of MST as the final stage of the motion processing pathway in the primate visual cortex. This chapter presents results from two psychophysics experiments with human subjects that explore how different stimulus features affect motion perception. The results suggest that the contributions of different areas along this motion processing pathway are flexible and depend on stimulus features.

Author contributions

B.W. and S.T. designed research. B.W. performed research. B.W. analyzed data. B.W. and S.T. wrote the manuscript. S.T. supervised the project.

Acknowledgements

We thank Ann-Kristin Kenkel and Ilona Vieten for help with data collection.

Abstract

Modern accounts of visual motion processing in the primate brain emphasize a hierarchy of different regions within the dorsal visual pathway, especially primary visual cortex (V1) and the middle temporal area (MT). However, recent studies have called the idea of a processing pipeline with fixed contributions to motion perception from each area into doubt. We address the idea in more detail in two experiments with human subjects that compare motion perception of two commonly used stimulus types: drifting sinusoidal gratings and random dot patterns (RDPs). Varying size and contrast levels, we confirm previous findings that increasing stimulus size impairs performance for high-contrast gratings. However, using our paradigm we did not observe a reversal of the size effect for low-contrast gratings. Furthermore, we did not find an effect of stimulus size on performance for RDPs with any contrast level. We conclude that the specific effects of stimulus size on performance observed previously do not easily generalize, limiting the fundamental insights about brain function that can be derived from this phenomenon.

7.1 Introduction

Our visual world is highly dynamic and the ability to perceive motion is essential for surviving and striving in an ever-changing world. In addition to its everyday relevance, motion processing is also an excellent and well-studied model system for visual neuroscience more generally. Motion is defined by only a few parameters (most importantly direction and speed) and a large number of studies have attempted to describe how these parameters are linked to fundamental neuronal properties (Block, 2005; Britten et al., 1992; Mauss et al., 2017). Furthermore, a limited number of well-defined regions in primate visual cortex are known to respond selectively to these parameters. The often implicitly assumed "standard model" of motion processing assigns key roles to the primary visual cortex (V1) and the middle temporal (MT) area (Albright, 1984; Born & Bradley, 2005; Maunsell & Van Essen, 1983c). Visual area 3 (V3) (Gegenfurtner, Kiper, & Levitt, 1997) and the medial superior temporal cortex (MST) (Mineault et al., 2012; Saito et al., 1986; Tanaka et al., 1986; Wild & Treue, 2021b) are further regions that play important roles for the processing of motion information. The importance of area MT in particular has been highlighted by a great number of studies suggesting that it is both necessary and sufficient for motion perception. Lesions in area MT lead to increased motion thresholds (Newsome & Paré, 1988; Rudolph & Pasternak, 1999) while micro-stimulating neurons in this area can bias a monkeys judgments of motion direction (Salzman et al., 1990, 1992), suggesting a direct link between MT activity and motion perception. This "standard model" has been questioned, however, in a recent study which suggests that alternate perceptual pathways bypassing area MT can be accessed under special circumstances (L. D. Liu & Pack, 2017). Monkeys were trained over several weeks to indicate the direction of a briefly presented Gabor grating of varying contrast levels. Recordings in MT confirmed that cells in that area were tuned to the direction of the gratings and that neuronal activity was correlated to behavior. Surprisingly, however, reversible inactivation of MT had little effect on behavioral performance, indicating that MT is not always necessary for motion perception and that other motion-sensitive areas (e.g., V1) can compensate for the disruption. Monkeys were then trained for several weeks on a similar task that used random dot patterns (RDPs) of varying coherence levels rather than gratings of varying contrast levels. Again, MT activity was correlated with behavioral performance, but this time MT inactivation did affect performance. When the grating experiment was repeated after the training with RDPs, MT inactivation, which had previously had very little effect on behavior in the grating task, led to a large performance impairment. The authors conclude that prolonged training with random dot

stimuli increased the contribution of area MT to perceptual decisions about motion, even for other types of motion stimuli. In addition to the effects that inactivating area MT had on motion perception, Liu & Pack (2017) also report a behavioral signature for a shift of perceptual processing between different brain regions: when testing the monkeys motion discrimination performance on high-contrast gratings of varying sizes, performance increased with stimulus size up to a certain point (the "optimal stimulus size"), before it declined for even bigger stimuli. This finding was first described by Tadin and colleagues (Tadin, Lappin, Gilroy, & Blake, 2003) and has been repeatedly replicated in humans (Betts et al., 2005, 2009; Serrano-Pedraza et al., 2011; Tadin & Lappin, 2005; Tadin et al., 2011; Yazdani et al., 2015, 2017) and monkeys (L. D. Liu et al., 2016). Originally it had been hypothesized that this effect is caused by the suppressive surround of receptive fields of individual MT neurons which is stimulated by gratings exceeding a certain size (Tadin et al., 2003). Physiological recordings and modeling, however, suggest that it is the surround suppression of several neurons and the correlation in activity between these neurons rather than suppression of individual neurons that causes the psychophysical suppression effect (L. D. Liu et al., 2016). In their more recent study, Liu and Pack (2017) found that the "optimal stimulus size" for gratings increased after training with RDPs, which they interpret as evidence that the prolonged training with the RDPs shifted perceptual processing to an area with larger receptive fields (e.g., from V1 to MT). Together, these results suggest that there is not a single processing pipeline for visual motion that works in the same way for all types of stimuli. Instead, the contributions of different motion-sensitive brain areas seem to be highly flexible and features of stimuli appear to play a role in determining which areas or local neuronal networks are recruited for the task at hand. This is in line with the "Reverse Hierarchy Theory" (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002) which suggests that neurons at different levels of the sensory processing pathways are recruited as needed to optimize perception and perceptual learning. To provide additional evidence for the existence of alternate motion processing pathways, we performed two experiments in which human subjects discriminated the motion of briefly presented grating and random dot stimuli of varying contrast levels and sizes. We find a clear effect of stimulus size for high contrast gratings, but inconsistent results for low contrast gratings. In contrast, we find no effect of size on RDPs at any contrast levels. We conclude that linking the effects of stimulus size on performance to something as fundamental as the surround structure of receptive fields in visual cortex has to be reconsidered.

7.2 Experiment 1

7.2.1 Methods

Participants

A total of 18 subjects were recruited for the experiment. Data from one subject had to be excluded due to technical problems during the experiment. Two further subjects were not invited back for the main experiment after an initial training session (see below for detailed description) because they described the task as extremely difficult and were not able to reliably report motion direction at stimulus durations below 166 ms. Of the remaining 15 subjects (9 female, 6 male; ages 20-34 years, mean age = 24.1 years), all were right-handed, had normal ($N = 8$) or corrected-to normal ($N = 7$) vision, and were naïve to the purpose of the experiment. All participants gave informed written consent prior to participating in the study. The study adhered to institutional guidelines for experiments with human subjects, was approved by the Ethics Committee of the Georg-Elias-Müller-Institute of Psychology, University of Göttingen (GEMI 17-06-06 171), and was in accordance with the principles of the Declaration of Helsinki.

Experimental setup

Subjects were seated in a dimly-lit room and viewed all stimuli on an LCD screen (SyncMaster 2233, Samsung) with a refresh rate of 120 Hz and a background luminance of 36 cd/m^2 at a distance of 57 cm. The experiment was controlled with the open-source software MWorks (`mworks-project.org`) running on an Apple MacPro computer. Subjects responded on a gamepad (Precision, Logitech). All stimuli were presented at the center of the screen (see the following section), and subjects were provided with a central fixation point between trials. Eye movements were not tracked.

Stimuli and procedure

Subjects were presented with either a horizontally drifting sinusoidal grating with a 2D Gaussian envelope ("Gabor patch") or a random dot pattern (RDP) and had to report whether a stimulus was moving to the left or to the right. Matching Tadin et al. (2003), the grating had a spatial frequency of 1 cycle per degree of visual angle. The radius of each grating was defined as twice the standard deviation of the Gaussian envelope (2σ). We used σ values of 0.4°, 1.0°, 1.6°, 2.2°, and 2.8°, resulting in gratings of radius 0.8°, 2.0°, 3.2°, 4.4°, and 5.6°. The phase of the grating at stimulus onset varied

randomly from trial to trial and moved with a speed of 3.5°/s to the left or the right. Three different contrast levels were presented in different blocks by changing the gratings' transparency. Michelson contrasts, calculated as $\frac{I_{max}-I_{min}}{I_{max}+I_{min}}$ (where I_{max} is the highest luminance (i.e., the brightest white) and I_{min} is the lowest luminance (i.e., the darkest black) of the grating) for what we call the "high", "intermediate", and "low" contrast condition were 99.2% [I_max: 70 cd/m^2, I_min: 0.3 cd/m^2], 8.3% [I_{max}: 39 cd/m^2, I_min: 33 cd/m^2], and 2.8%. [I_{max}: 37 cd/m^2, I_min: 35 cd/m2]. We chose a relatively low contrast value (8.3%) for "intermediate" because Tadin et al. (2003) showed the effect of size change to lie between 2.8% and 11%, with little qualitative difference between higher contrast levels. Random dot patterns had radii of 1°, 2.5°, 4°, 5.5°, and 7° (which, perceptually, resembled the size of our gratings) and a dot density of 4 dots/deg^2 (i.e., between 12 and 615 dots). All dots were placed randomly within the aperture and moved either to the left or to the right with a speed of 2 °/s. Again, three different contrast levels (I_{max} and I_{min} now being the luminance of the dots and of the background, respectively) were used in different blocks: 32% [I_{max}: 70 cd/m^2, $I_m in$: 36 cd/m^2], 13.3% [I_{max}: 47cd/m^2, I_{min}: 36 cd/m^2], and 5.3% [I_{max}: 40 cd/m^2, $I_m in$: 36 cd/m^2], termed "high", "intermediate", and "low" contrast, respectively. Our two stimulus types differed in a number of parameters: speed (3.5 °/s for gratings, 2 °/s for RDPs; (Tadin et al. (2003): gratings: 2 °/s; RDPs: 8 °/s)) was chosen based on pilot studies so as to achieve an intermediate performance for each stimulus type at an intermediate stimulus size. Contrast levels differed, to keep a fixed background luminance across the two stimuli and because contrast depends on the difference in luminance within the stimulus for gratings, but on the difference between stimulus luminance and background luminance for the RDPs. Finally, the exact size of the Gabors cannot be determined, as their contrast fades out with increasing distance from the stimulus center and perceived size depends on the standard deviation and peak-contrast of the grating (Fredericksen, Bex, & Verstraten, 1998). However, we are not comparing absolute performance between stimulus types for a given contrast level and size, but rather investigating the effects of varying these parameters on the performance for a given stimulus type. In other words, we are making claims of the type "Increasing stimulus size effects performance for gratings, but not RDPs", rather than "A grating of size X and contrast Y is perceived better than a RDP of size X and contrast Y". Subjects were asked to foveate the center of the screen and started each trial with a button press. The fixation dot then disappeared and the stimulus was presented for a brief duration that had been adjusted for each subject individually based on a training session (see below for further information). Subsequently, the stimulus was masked for 220 ms with an RDP of radius 15 that had 2,000 black and 2,000 white

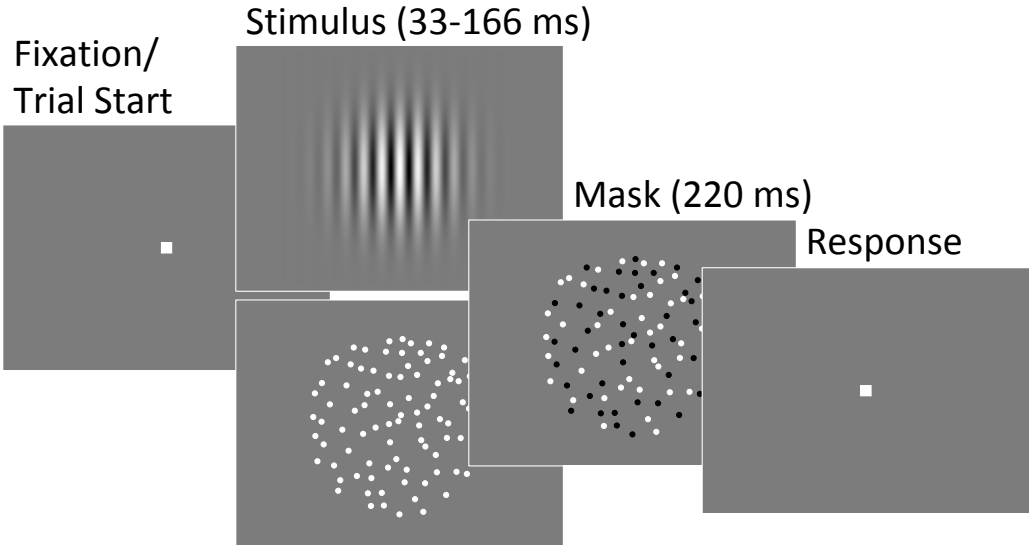

Figure 7.1: Example trial sequence in Experiment 1: subjects initiated each trial with a button press. Either a drifting grating or a random dot pattern moving to the left or to the right was then briefly presented for a duration that was adjusted for each subject individually to ensure a significant but not perfect performance of about 75% for the intermediate stimulus size. A large, high-contrast random dot pattern with black and white dots moving in random directions was used as a mask. Note that these are not screenshots of the actual stimuli and that this figure has been optimized (e.g., in terms the grating's Gaussian envelope or the RDPs dot-size and number of dots) to illustrate the sequence of events.

dots (density: 5.7 dots/deg^2), all moving in random directions at 15 °/s. After the mask disappeared and the fixation point reappeared subjects reported their perceived direction (left vs. right) by pressing the corresponding trigger button on the gamepad (see Fig. 7.1).

The experiment comprised 18 blocks of 100 trials each: For each stimulus type (grating and RDP) and every contrast level (high, intermediate, low), subjects completed 3 blocks (300 trials). Within a block, every stimulus size was presented 20 times in random order. Thus, each subject saw every possible combination of a specific stimulus type, contrast, and size 60 times. Blocks with the same stimulus type and blocks with the same contrast level were presented after one another: the three blocks with a high-contrast grating were followed by three blocks of the intermediate-contrast and three blocks of the low-contrast gratings and then a total of nine blocks of high-, intermediate-,

and low-contrast RDPs. To familiarize subjects with the task, they underwent a training session that took place 1-7 days before the actual test session. Each subject practiced with at least 1,000 trials, using gratings and RDPs of varying sizes and contrast levels (though not necessarily the exact values that were used during the experiment). The presentation duration of the stimulus was varied during the training session, starting with long durations until subjects had understood the task well. Towards the end of the training session, one stimulus duration for gratings and one for RDPs of the highest-contrast level was determined manually for which the subjects achieved approximately 75% correct responses for a stimulus of the intermediate size of the 5 sizes used in the main experiment. Due to technical problems, stimuli were randomly presented for one frame (∼8.3 ms) longer or shorter than specified on some trials. As this error occurred randomly across the different sizes it should have no systematic influence on the results. Indeed, excluding trials in which the stimulus duration differed by more than 8 ms from the most common stimulus duration for a given subject and condition (i.e., the one that had originally been specified) did not qualitatively change the results or alter the conclusions of our study.

Data analysis

Relevant data for the analysis (subject ID, stimulus type and contrast level, stimulus size, motion direction, button press, response correctness, stimulus duration) were extracted from the MWorks files using custom Matlab (MathWorks, Natick, MA) scripts. All further analyses were carried out in Python.For each combination of stimulus type, contrast level, and size we calculated the percentage of correct responses (out of 60 presentations) for each subject as a measure of how well they were able to perceive the stimulus direction. To investigate whether performance was affected by stimulus size and how this size-effect was influenced by contrast, we employed two analysis steps: First, we wanted to determine whether a subject's performance was influenced by stimulus size for a given combination of stimulus type and contrast level. For this purpose, we calculated Pearson's correlation coefficient between stimulus size and performance for each subject for each of the six combinations of stimulus type and contrast level. For each correlation coefficient we calculated a 95% confidence interval (CI) using a bootstrapping method in which we drew 2000 random samples with replacement, calculated the correlation coefficient for each sample, sorted the 2000 coefficients, and took the 50th and the 1950th value as the borders of the CI. We considered a subject to show a "negative size effect" (i.e., performance decreases with increasing stimulus size) if the entire 95% CI was smaller than 0 and a "posi-

tive size effect" (i.e., performance increases with increasing stimulus size) if the entire 95% CI was larger than 0. If a confidence interval crossed 0, we considered the subject not to show a significant size effect for that condition.

Second, we calculated a three-way repeated-measures analysis of variance (rmANOVA), with "stimulus type" (grating or RDP), "contrast" ("high", "intermediate", or "low") and "stimulus size" (5 levels) as within-subject factors.

7.2.2 Results

We tested how stimulus size and contrast affect the perception of direction of moving sinusoidal gratings ("Gabor patches") compared to random dot patterns (RDPs) as measured by the percentage of correct choices in a direction discrimination task. Stimulus durations were adjusted for each subject and each stimulus type (grating or RDP) individually. But the average stimulus

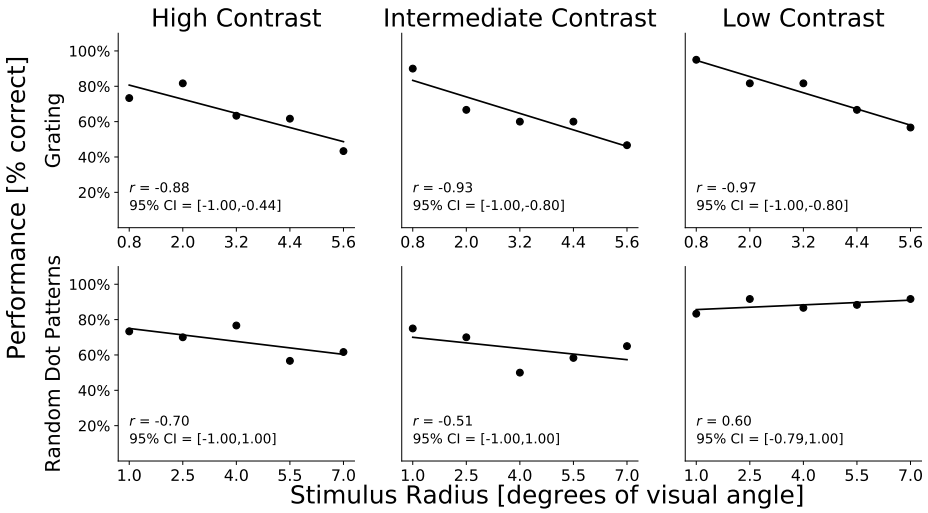

Figure 7.2: Effect of stimulus size on direction discrimination performance for an example subject in six different conditions: the top row shows data for the drifting grating, the bottom row for the random dot pattern. The three columns show data for three different contrast levels of the respective stimulus. This subject fulfilled our criteria for a negative size effect (negative correlation with a 95% confidence interval that lies entirely below 0) for all contrast levels for the grating, but not for any contrast level for the random dot pattern. Pearsons correlation coefficient and the 95% confidence interval are shown in each graph.

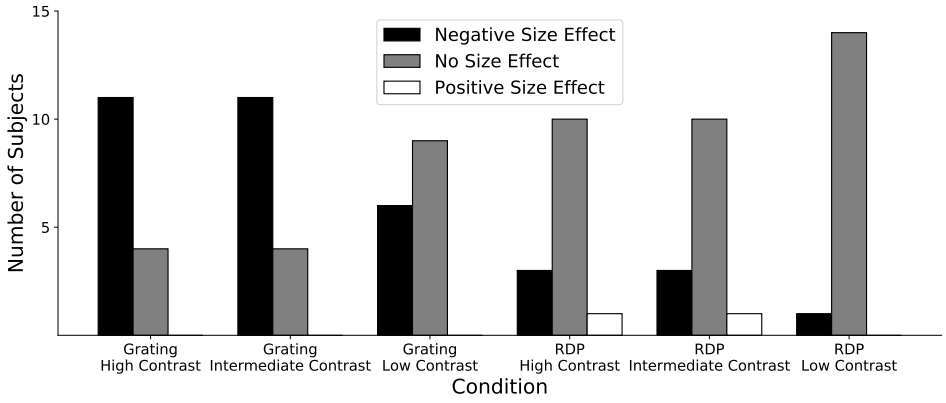

Figure 7.3: The number of subjects that showed a negative size effect (black bars), no size effect (gray bars) or a positive size effect (white bars) in each of the six conditions. See Methods for a definition of size effect.

duration across subjects did not differ significantly between gratings (M = 77.2 ms, SD = 33.85 ms) and RDPs (M = 73.3 ms, SD = 40.73 ms) (paired t-test: $t(14) = 0.76, p = .46$). We had manually determined a duration for each stimulus type for which observers should achieve a performance level of approximately 75% for high-contrast stimuli of the intermediate size. Indeed, one-sample t-tests confirmed that performance levels across all subjects did not differ significantly from 75% for the 4° high-contrast RDP ($t(14) = -0.15, p = .88$) or the 3.2° high-contrast grating ($t(14) = 0.17, p = .87$).

Figure 7.2 plots the results of a representative subject. For the grating (top row) the subject meets our criterion for a negative size effect for all three contrast levels: the correlation coefficient between stimulus size and performance is negative and the 95% confidence interval is entirely below 0. For the random dot patterns (bottom row) the subject shows no size effect for either contrast level according to our criteria: for all three cases, the 95% confidence interval crosses 0. Note also that the absolute values of the correlation coefficients are much lower than for the gratings.

Across our 15 subjects, at least six show a negative size effect for each of the three contrast levels of the grating and none shows a positive size effect. For RDPs, on the other hand, less than a third show a negative size effect and most subjects show no size effect at all (Fig. 7.3).

There is a significant main effects of "stimulus type", based on a three-way repeated measures ANOVA ($F(1, 14) = 11.89, p < .01$) and "stimulus size" ($F(4, 56) = 30.87, p < .001$), but no main effect of "contrast" ($F(2, 28) = 1.30, p = .29$). The interactions between "stimulus type" and "stimulus size" ($F(4, 56) = 20.81, p < .01$) and between "contrast" and "stimulus size"

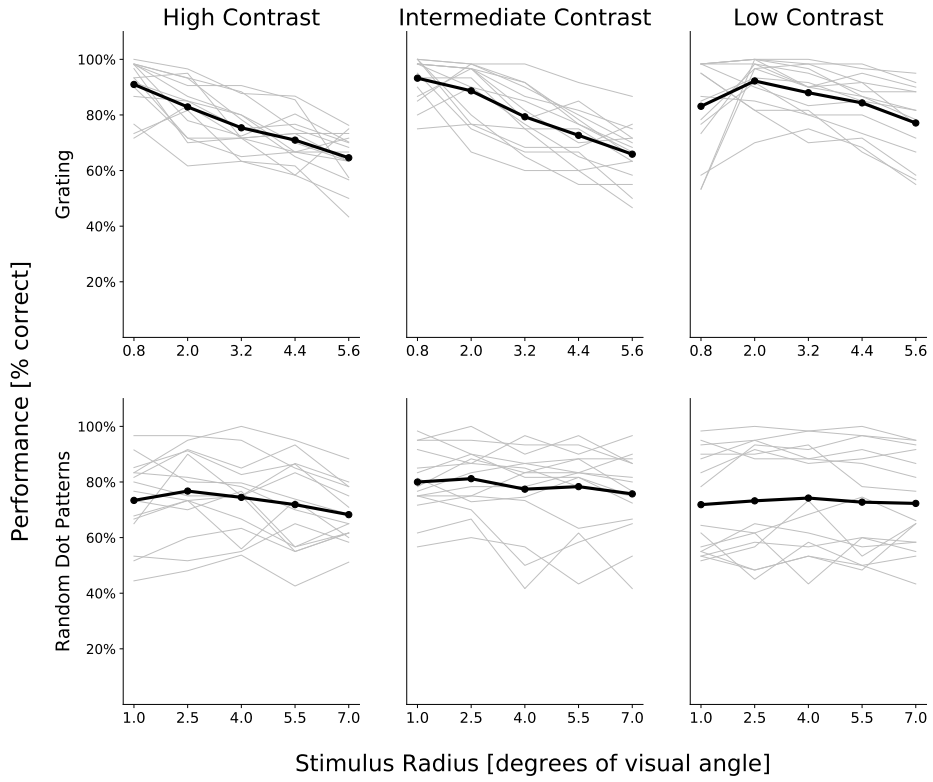

Figure 7.4: Effect of stimulus size on direction discrimination performance across all 15 subjects. The large black dots and lines show the mean across subjects, the thin gray lines show data from individual subjects. All other information is depicted as in Fig. 7.2.

($F(8, 112) = 6.95, p < .01$), as well as the three-way interaction between "stimulus type", "contrast", and "stimulus size" ($F(8, 112) = 3.47, p < .01$) were significant, but not the interaction between "stimulus type" and "contrast" ($F(2, 28) = 1.64, p = .20$). In other words, the effect of stimulus size depends on the stimulus type and on contrast. This can be clearly seen in the averaged data of all subjects for all six conditions, which show impaired performance with increasing stimulus size across all contrast levels for gratings (Fig. 7.4, top row), but not random dot patterns (Fig 7.4, bottom row).

7.2.3 Interim discussion

We observed two surprising differences compared to previous studies: first, while we were able to replicate the detrimental effect of increasing stimulus size on performance for high-contrast gratings, we observed no such effect for random dot patterns. Second, we could only replicate the finding that performance decreases with increasing stimulus size for high-contrast gratings, but we did not observe the reverse effect for low contrast gratings that had previously been described (Tadin et al., 2003). There are a few, potentially important issues that could explain the difference between our results and those reported in the literature: One major difference of our study, compared to previous experiments (Tadin et al., 2003; Yazdani et al., 2015; Serrano-Pedraza et al., 2011), is the use of a mask right after the moving stimuli. Visual persistence, i.e., the extended perception of a visual stimulus after the stimulus has been terminated, has been shown to be inversely related to stimulus duration and luminance (Bowen et al., 1974). Therefore, it is likely that visual persistence varied across the different contrast-levels, stimulus sizes, and stimulus durations of other and our studies. Second, we presented the different stimulus types and contrast levels in a blocked design with the order of blocks being identical across subjects. Thus, differences between gratings and RDPs (and potentially between different contrast levels) could be due to training effects. Third, our two stimulus types (gratings and RDPs) differed in a number of features, such as speed and the overall luminance (which is kept at the level of the background for sinusoidal gratings, but increases above the background for RDPs with white dots).

To address all these issues, we conducted a second experiment in which we presented all combinations of stimulus type, contrast, and size randomly, used RDPs that consisted of black a white dots, to ensure an overall stable luminance level, and used a mask only on half of all trials.

7.3 Experiment 2

7.3.1 Methods

Participants

Our aim is to collect data from 16 participants. Previous studies that addressed the same research questions had between 4 and 6 participants (Tadin et al., 2003; Serrano-Pedraza et al., 2011; Tadin et al., 2019). However, we are aiming to investigate effects that are likely to be smaller (such as differences between gratings and random dot patterns) and therefore chose a 3-fold larger

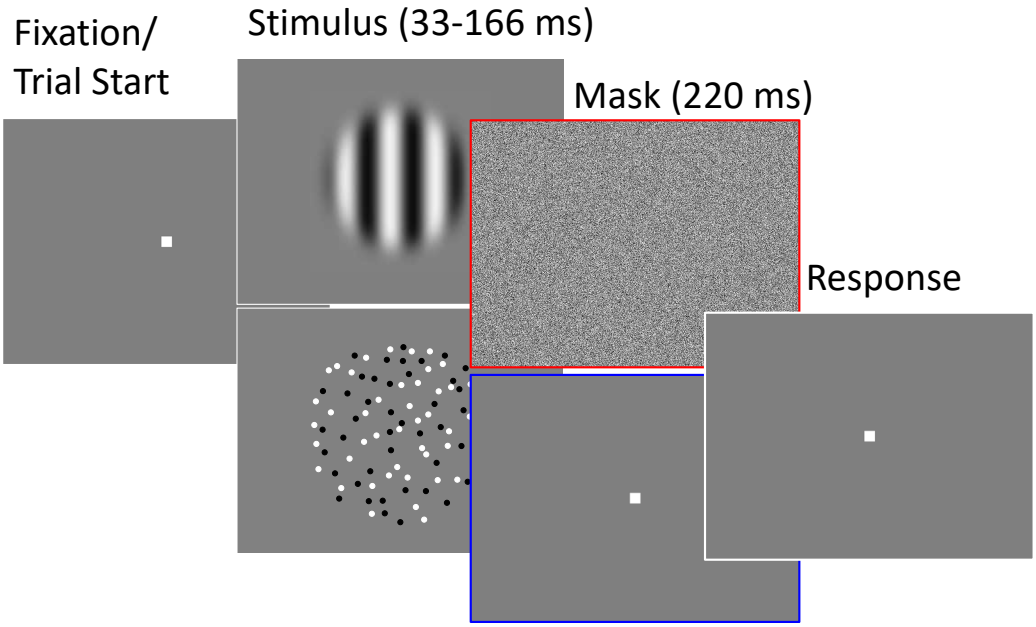

Figure 7.5: Example trial sequence in Experiment 2: the two main differences to Experiment 1 are that the mask was replaced with flickering white noise and that the mask was only presented on 50% of trials.

sample size, increasing the sensitivity for small effect sizes. Currently datasets from 8 subjects (4 female and 4 male) have successfully been recorded. Six subjects were right-handed, the mean age was 25.6 years (range: 22-31) and all had normal ($N = 6$) or corrected-to normal ($N = 2$) vision. Two subjects were actively involved in the design of the experiment, two further subjects were familiar with the basic research question behind the experiment, and the remaining four subjects were naïve to the purpose of the experiment. The six participants not involved in the design of the experiment gave informed written consent prior to participating in the study. The study adhered to institutional guidelines for experiments with human subjects, was approved by the Ethics Committee of the Georg-Elias-Müller-Institute of Psychology, University of Göttingen (GEMI 17-06-06 171), and was in accordance with the principles of the Declaration of Helsinki.

Experimental setup

The setup was identical to Experiment 1, except that eye movements were tracked and trials on which the subject's gaze left a $3° \times 3°$ square window around the fixation point were aborted and repeated later in the experiment.

Stimuli and procedure

The stimuli and procedure were the same as in Experiment 1 expect for the following changes:

- The RDP consisted of 50% white and 50% black dots to keep the average luminance across the screen stable.
- The grating and the RDP moved with the same speed (2.5 °/s).
- The contrast levels for both stimuli were the same: "high": 99% (I_{max}: 78 cd/m^2, I_min: 0 cd/m^2); "intermediate": 8% (I_{max}: 19 cd/m^2, I_min: 16 cd/m^2); "low": 3% (I_{max}: 18 cd/m^2, I_min: 17 cd/m^2).
- The mask was presented on only half of the trials and instead of a random dot pattern, flickering white noise across the entire screen was used as a mask (see Fig.7.5).
- Every subjects completed 12 blocks of 240 trials (2880 trials in total) so that every combination of 2 stimulus types, 3 contrast levels, 5 stimulus sizes, and 2 masking conditions (with and without mask) was presented 48 times.
- Most importantly, the different conditions were presented in random order, rather than in blocks as had been the case in Experiment 1.

Subjects still underwent a training session 1-7 days before the actual test session and completed at least 1000 trials of training. In contrast to Experiment 1, we determined a single stimulus duration for each subject that was used for all conditions.

Data analysis

Data analysis was similar to Experiment 1: For each combination of stimulus type, contrast level, size, and masking condition we calculated the percentage of correct responses (out of 48 presentations) for each subject as a measure of how well they were able to perceive the stimulus direction.

We again defined a "size effect" (same definition as in Experiment 1) for every combination of stimulus type, contrast, and masking conditions, and calculated a four-way repeated-measures analysis of variance (rmANOVA), with "stimulus type" (grating or RDP), "contrast" ("high", "intermediate", or "low"), "masking condition" (with or without mask) and "stimulus size" (5 levels) as within-subject factors.

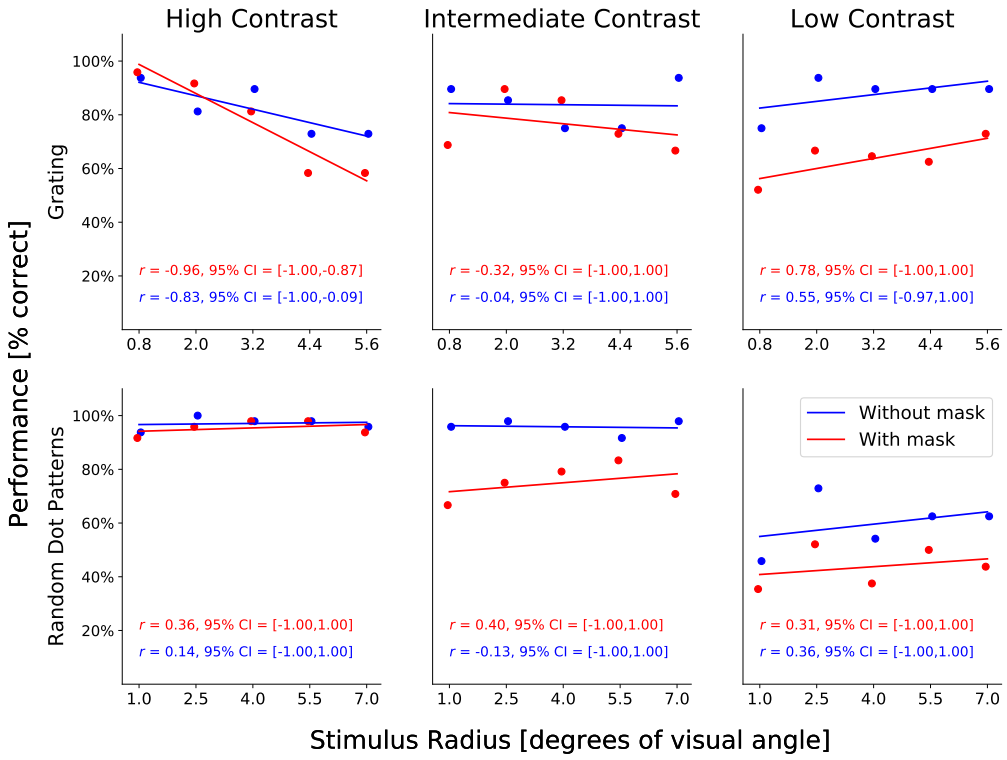

Figure 7.6: Effect of stimulus size on direction discrimination performance for an example subject in 12 different conditions: the top row shows data for the drifting grating, the bottom row for the random dot pattern. The three columns show data for three different contrast levels of the respective stimulus. The color indicates the performance on trial with (red) or without (blue) a mask. This subject fulfilled our criteria for a negative size effect (negative correlation with a 95% confidence interval that lies entirely below 0) only for the high contrast grating condition in both masking conditions.

7.3.2 Results

Because we used a single stimulus duration for both gratings and RDPs in Experiment 2, it was more difficult to find a value that ensures an interpretable performance without any floor or ceiling effects across all conditions. Whereas for gratings, the average performance across subjects for the high-contrast stimulus of the intermediate size in the "with mask" condition did not differ significantly from 75% ($t(14) = 1.87, p = .10$), for RDPs performance was significantly better than 75% ($t(14) = 40.39, p < .001$).

Figure 7.6 plots the results of a representative subject. The plot is structured the same way as Fig. 7.2, but the two masking conditions are

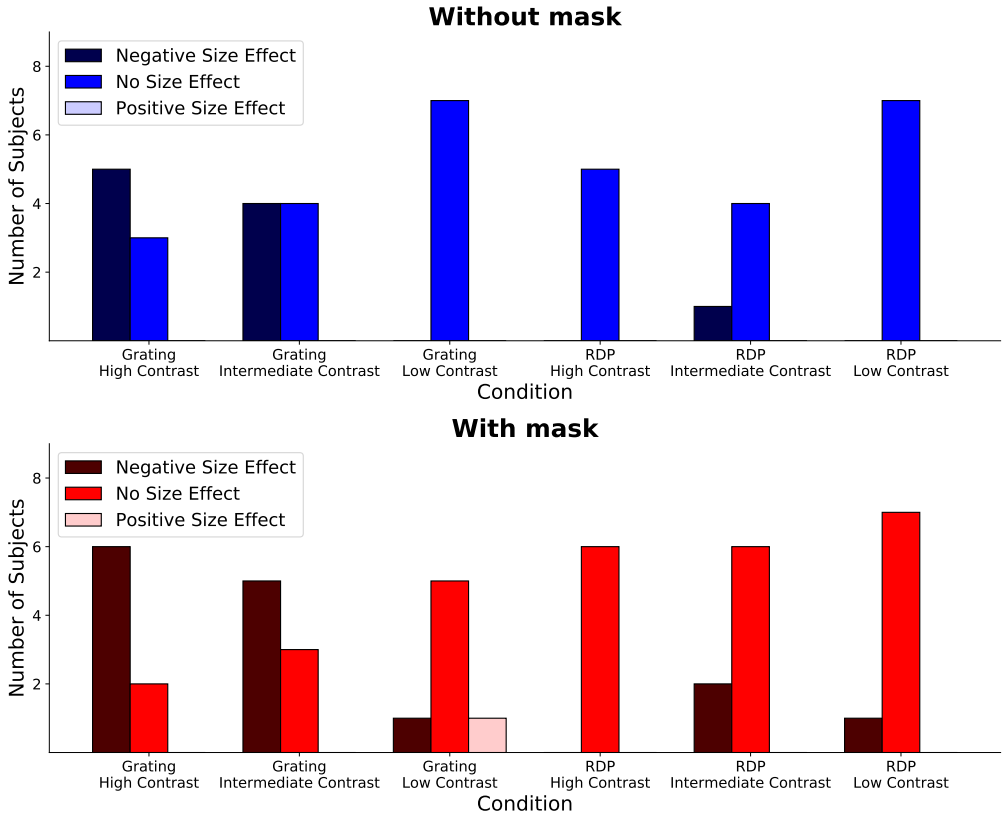

Figure 7.7: The number of subjects that showed a negative size effect (black bars), no size effect (gray bars) or a positive size effect (white bars) in each of the six conditions in trials without (top plot) or with (bottom plot) a mask.

shown in red (with mask) and blue (without mask). For the grating (top row) the subject meets our criterion for a negative size effect only for the high-contrast condition (top left plot), in both, the "with mask" (red) and "without mask" (blue) conditions. For the other combinations of stimulus type and contrast (remaining 5 plots) the subject shows no size effect for either masking condition: for all ten cases, the 95% confidence interval crosses 0.

The data from all 8 subjects confirms that a negative size effect was prevalent for the high contrast condition and present in some subjects for the intermediate contrast gratings, but that there were no consistent effects of stimulus size in any of the other conditions (Fig. 7.7). Interestingly, the presence or absence of a mask seemed to have little to no effect (compare top and bottom panel of Fig. 7.7)

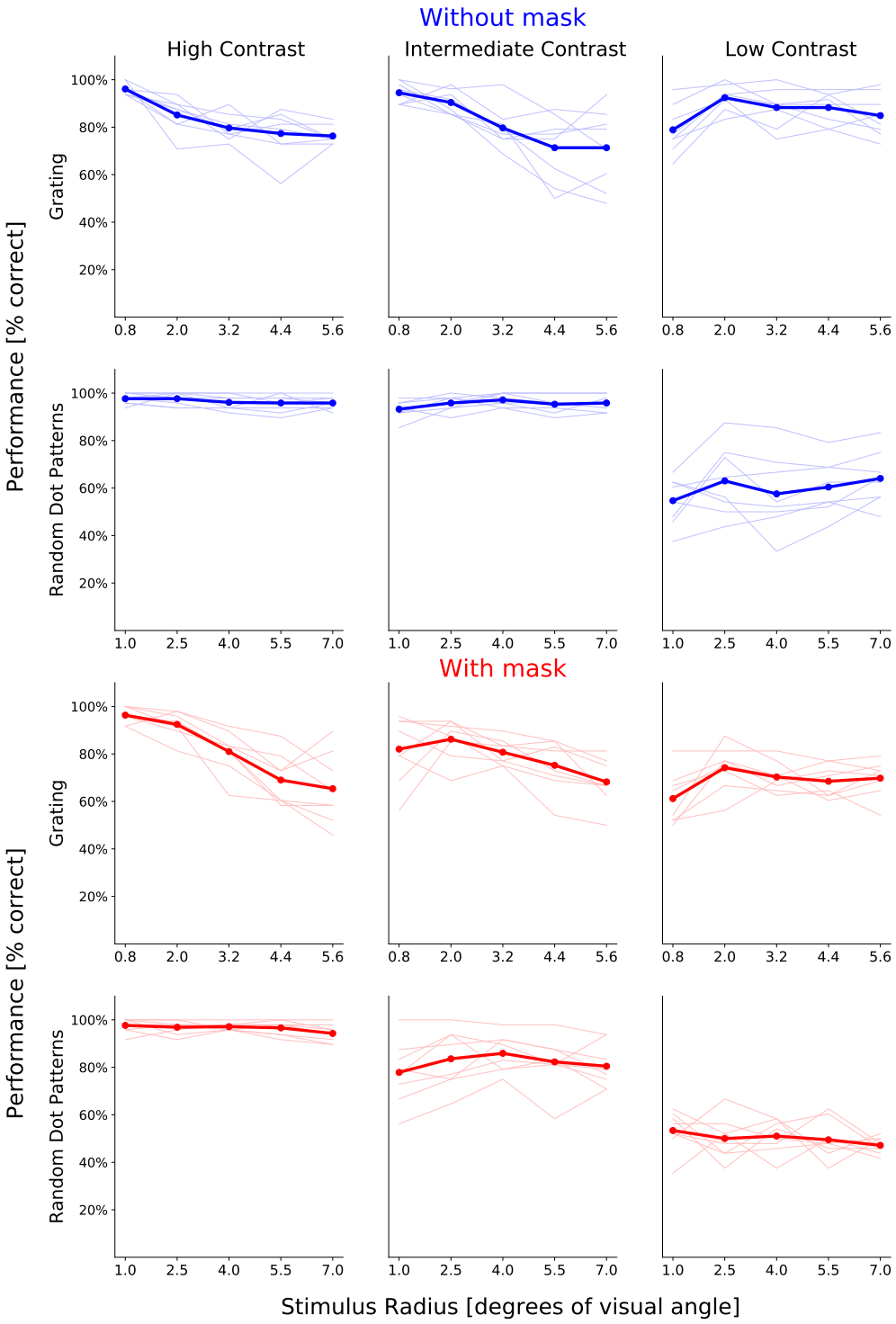

Figure 7.8: See next page for legend.

Figure 7.8: *(continued)* Effect of stimulus size on direction discrimination performance across all 8 subjects. The large dots and thick lines show the mean across subjects, the thin lines show data from individual subjects. First two rows (red) show data from trials without a mask, bottom two rows (blue) show data from trials with mask. All other information is depicted as in Fig. 7.2.

The average performance across all subjects for all conditions is shown in Fig. 7.8. A four-way repeated measures ANOVA revealed main effects of "size" ($F(4, 28) = 11.99, p < .001$), "contrast" ($F(2, 14) = 122.96, p < .001$), and "masking condition" ($F(1, 7) = 32.10, p < .001$), but not "stimulus type" ($F(1, 7) = 0.02, p = .89$). Most importantly, the four-way interaction "stimulus type" × "contrast" × "masking condition" × "size" was significant ($F(8, 56) = 3.64, p < .01$), indicating that each of the main effects depends on the other variables. To explore these relations in more detail, we performed 4 2-way repeated measures ANOVAS with factors "size" and "contrast" for each combination of stimulus type and masking condition with a Bonferroni corrected significance level of $0.05/4 = 0.0125$. For the sake of readability, we only report significant main effects and interactions.

For gratings without a mask, we found a significant main effect of "size" ($F(4, 28) = 11.47, p < .001$) and a significant interaction between "size" and "contrast" ($F(8, 56) = 9.24, p < .001$). The results for gratings with a mask were similar, except that the main effect of contrast was also significant ($F(2, 14) = 10.18p < .01$) in addition to the main effect of "size" ($F(4, 28) = 11.96, p < .001$) and the interaction ($F(8, 56) = 10.91, p < .001$). In both cases (with and without mask), post-hoc Tukey HSD tests showed that there were no significant differences (at a significance level of 0.0125) in performance between different stimulus sizes in the low contrast condition[1] (see the top right panels in the top (blue) and bottom (red) half of Fig. 7.8), but in the high and intermediate contrast performance was better for the smallest than for the two largest sizes.

For RDPs, we found only a significant main effect of "contrast" (without mask: $F(2, 14) = 77.50, p < .001$; with mask: $F(2, 14) = 142.00, p < .001$), but no effects of "size" an no interaction, suggesting that stimulus size does not modulate performance. Post-hoc Tukey HSD tests showed that in trials with a mask performance was better in the "high" than in the "intermediate" or "low" contrast condition and better in the "intermediate" than in the "low"

[1]The difference between the smallest and the second-smallest stimulus is significant at a significance level of 0.05, but not at the significance level of 0.0125 that we use here to account for multiple comparisons

contrast condition, whereas in trial without a mask, there was no difference between "high" and "intermediate" contrast levels, but for both performance was better than for "low" contrast (all $p < .01$).

7.3.3 Interim discussion

Experiment 2 supported the three main findings of Experiment 1.

First, for high contrast gratings, increasing stimulus size leads to decreases in performance, in agreement with previous studies (e.g., Tadin et al., 2003; Yazdani et al., 2015; Serrano-Pedraza et al., 2011).

Second, this effect disappears for low contrast gratings. We did not find evidence for a reversal of the effect, as originally reported by Tadin and colleagues (2003). In Experiment 2, there seemed to be a trend for an increase in performance from the smallest to the second-smallest stimulus size, but more statistical power (i.e., more subjects) is necessary to determine whether this effect is significant. As the statistical power needed to detect an effect depend on the effect size, we can conclude that a potential increase in performance, which our experiment was unable to detect, would be much smaller than the detrimental effect of size on performance for high contrast gratings.

Third, we found no evidence for an effect of stimulus size on performance in a motion discrimination task with random dot patterns (see General Discussion).

The presence of absence of a mask appeared to have little effect other than to decrease performance for intermediate contrast RDPs. The major difference in the results between Experiment 1 and the trials of Experiment 2 with a mask (i.e., Figures 7.4 and bottom half of 7.8) is the effect of contrast on performance with RDPs: whereas performance is very stable across all three contrast levels in Experiment 1 (blocked design), it decreases with decreasing contrast in Experiment 2 (randomized design). Note however, that in Experiment 1 for low contrast RDPs (bottom right panel of Fig. 7.4), roughly half of the subjects show consistent performance above the average and the other half show consistent performance below average. This could indicate that subjects used different strategies to deal with this difficult condition, but further research is needed to explore this in more detail.

7.4 General discussion

Drifting sinusoidal gratings and random dot patterns are popular and powerful stimuli to investigate motion perception in psychophysics (Tadin et al., 2003; Adelson & Movshon, 1982; Busse et al., 2008; Curran et al., 2019; Morrone et al., 1995; Treue et al., 1993), functional imaging (Castelo-Branco et al., 2002; Mikellidou et al., 2018; Sack et al., 2006), and electrophysiology (Britten et al., 1992; Pack, Berezovskii, & Born, 2001; Snowden et al., 1992; Treue & Martinez-Trujillo, 1999; Treue & Maunsell, 1996). The popularity of these stimuli is due to the fact that they can be carefully created to isolate features of interest (e.g., direction of motion), with little confounding by other features (e.g., color, or emotional associations) (see Rust & Movshon, 2005, for a discussion of the advantages of artificial, synthetic stimuli). However, little research has been dedicated to investigating whether there are relevant differences in how these two stimulus types are processed and perceived.

Our study compared the effects of size and contrast on how well the direction of motion of those two stimulus types can be discriminated. We observed two surprising differences compared to previous studies: first, while we were able to replicate the effect of stimulus size on performance for high-contrast gratings, we observed no such effect for random dot patterns. Second, we could only replicate the finding that performance decreases with increasing stimulus size for high-contrast gratings, but we did not observe the reverse effect for low contrast gratings that had previously been described (Tadin et al., 2003).

Differences between gratings and random dot patterns

We found that performance systematically decreased for increasing stimulus size for gratings, but not for RDPs (Fig. 7.4 and 7.8). In the original study that first reported the effect of stimulus size on motion discrimination (Tadin et al., 2003), a random-dot stimulus consisting of light and dark pixels, half of which shifted in one direction while the other half was randomly regenerated, showed the same size effect and the same contrast dependence as gratings. This differs from our "classical" RDP with isolated dots that all move in the same direction. A recent study replicated the effect for random-texture patterns consisting of light and dark elements (Tadin et al., 2019), which combine the individual elements of RDPs and the dense structure of gratings. Our experiment, on the other hand, found no effect of stimulus size for 100% coherent RDPs with a dot density of 4 dots/deg^2. Together, these results indicate that coherence, i.e., the percentage of dots moving in the same direction, as well as overall dot density are important factors that determine

whether size has an effect on the perceivability of RDPs. These differences suggest that previous findings do not easily generalize, but might depend on a limited set of stimulus parameters.

Only a few other studies have systematically manipulated the size and contrast of random dot patterns in a simple discrimination task. Morrone and colleagues (1995, Burr et al., 1998) varied the contrast and the overall area of RDPs by manipulating the number of visible sectors, rather than the radius. Their results support a two-stage model of motion processing: an initial stage consisting of contrast-sensitive local motion detectors and a second stage that integrates across the detectors and is tuned for more complex motion patterns. It should be noted that these studies used not only translational but also radial (expansion and contraction) and rotational motion patterns, which are presumably processed in area MST (Graziano et al., 1994; Wild & Treue, 2021b), an area in which center-surround antagonism is much less prominent than in MT or V1 (Tanaka et al., 1986; Lagae et al., 1994). Nevertheless, the consequences of the hypothesis that different areas along the motion processing pathway are particularly well suited to encode different stimulus features such as contrast, density, or direction deserve some further investigation.

It is a broadly accepted idea that early motion-sensitive areas of the primate cortex (e.g., V1) are predominantly concerned with local motion information, whereas later areas (e.g., MT) respond preferentially to global motion information (Curran et al., 2019; Castelo-Branco et al., 2002; Movshon et al., 1985). This increase in spatial integration can be explained, at least partly, by the increase in receptive field size along the processing pathway from V1 to MT (e.g., Mikami, Newsome, & Wurtz, 1986). Gratings require little spatial integration as all the available motion and orientation information can be extracted from any small portion of the stimulus. For RDPs, on the other hand, a small part of the stimulus might not contain any dots (depending on the dot density) or have a low signal-to-noise ratio (depending on coherence) so that spatial integration is necessary to perceive a well-defined motion direction. Thus, a plausible (but presumably oversimplistic) hypothesis would be that perception of gratings relies primarily on activity in V1 whereas that of RDPs requires activity in area MT. As Neri & Levi (2006) have pointed out, different stages along the visual hierarchy can impose different constraints upon further processing which sometimes can be observed in the final behavioral output. It is in that sense that we suggest V1 and MT to have different roles in the processing of RDP and grating motion. The hypothesis that different areas of the dorsal visual pathway vary in their contribution to the perception of different stimuli receives further support from a recent study that investigated the effects of stimulus complexity on the specificity of

perceptual learning (Bakhtiari, Awada, & Pack, 2020). The authors found that training with RDPs (which are presumably processed primarily in area MT) reduced the spatial specificity of a direction discrimination task with drifting gratings and training with optic flow patterns (which are presumably processed primarily in area MST) reduced the spatial specificity of a direction discrimination task with translational RDPs. These results could be well explained by a computational "reweighting" model in which the readout from areas MT and MST shifted depending on training history (Bakhtiari et al., 2020).

No effect of contrast on the size-effect in gratings

Another result from Tadin et al.'s original study (2003, see also Yazdani et al., 2015; Serrano-Pedraza et al., 2011; L. D. Liu et al., 2016) that we did not observe is the effect of contrast. These previous studies reported that performance improves with increasing stimulus size for low-contrast gratings (presumably because of spatial summation) but deteriorates with increasing stimulus size for high-contrast gratings (presumably because of surround inhibition). In contrast, our data show a negative relationship between stimulus size and performance for high and intermediate contrast gratings and no effect for low contrast gratings. There was a trend for the average performance across subjects to increase slightly from the smallest to the second-smallest stimulus size (Fig. 7.8). But even this slight hint at an increase in performance for low contrast gratings might simply be due to the fact that a very small low contrast grating is difficult to see, and not be a specific effect on motion perception. We did not observe a positive size effect for any subject for any of the contrast-levels, as one would have expected based on Tadin et al.'s (2003) results. There are a few, potentially important differences between the previous studies (Betts et al., 2005; Serrano-Pedraza et al., 2011; Tadin & Lappin, 2005; Yazdani et al., 2017) and ours: first, whereas the previous studies measured the threshold exposure duration that was necessary for subjects to reach a certain performance level, we measured performance with a stimulus duration that was fixed for each subject. Thus, the results are difficult to compare directly without knowing the evidence accumulation process that is underlying the decision. The previous studies report that subjects needed a certain amount of time to reach a specified performance level for a given stimulus size. Assume in our experiment subjects have 20 ms more time than the threshold duration determined by Tadin and colleagues, but we find, unexpectedly, that their performance is not better than the level that the staircase in Tadin et al. (2003) was aiming for. This could mean that sensory evidence is accumulated very quickly up to a certain

level, but then it takes proportionally more time to improve upon this level. Unfortunately, it is impossible to investigate this hypothesis in more detail with the available data, as the exact stimulus sizes and durations vary between the two studies.

Second, there were a number of technical differences between the studies, such as in background luminance or the use of a temporal Gaussian window in most of the previous studies (Tadin et al., 2003; Tadin & Lappin, 2005; Serrano-Pedraza et al., 2011), compared to a rectangular window in our study. Differences in the temporal window can change the energy in different temporal frequencies, which is well-known to affect perception (Robson, 1966). However, as mentioned before, if this is a decisive factor, it calls into question the generalizability of the effect, instead suggesting that it depends on very specific parameters.

Effects of lesions and training

Our current experiment does not address the question what the visual system can do when forced to compensate for failures of certain subsystems or after prolonged training. In our experiment, all subjects did at least 1,000 trials of training to fully familiarize them with the task, but it is unclear whether this had any effect on the results we obtained. Two recent studies from Pack and colleagues have demonstrated a high degree of plasticity in the motion processing system: monkeys that originally showed little impairment for the grating discrimination task after MT inactivation, were clearly impaired in the same task after prolonged training with RDPs. In other words, training with RDPs changed the role that MT plays in motion perception, even for stimuli that the brain could previously perceive without MT (L. D. Liu & Pack, 2017). Training human observers with complex motion stimuli led to perceptual learning effects that are in line with a shift of the perceptual readout to higher brain regions in the dorsal visual pathway (such as from V1 to MT or from MT to MST) (Bakhtiari et al., 2020). As mentioned above, a possible interpretation of these results is that motion perception does not rely on a single, fixed processing pipeline, but rather on a flexible network where alternative processing paths vary their contributions depending on stimulus features and perceptual experience. It is then all the more surprising that such a network that demonstrates high flexibility in some cases is not able to compensate for the loss of MT after prolonged exposure to RDPs. The exact mechanisms that lead to this loss of flexibility and adaptability will have to be the subject of additional research. The currently available results suggest that the ability to flexibly shift between different processing modules is ironically not that flexible and that we should be cautious that

the effects of stimulus size on performance are not as universal and clear-cut as previously suggested.

Chapter 8

General discussion

8.1 Summary and interpretation

The overarching goal of this thesis was to gain new insights into how motion – especially complex motion with more than one linear direction – is represented in the primate brain.

Chapters 2 and 3 highlighted the medial superior temporal area (MST) of the extrastriate cortex as one key region for the processing of visual motion. An extensive review of the scientific literature showed that MST is a high-level visual area that encodes information about optic flow patterns and combines this information with vestibular signals to represent self-motion and object motion. Furthermore, MST plays an important role in cognitive processes, such as attention and working memory, and can thus be seen as a gateway between sensation and cognition.

Chapters 4 and 5 described a series of experiments that investigated the responses of MST neurons to visual motion in more detail. Automated mapping and tuning procedures confirmed that neurons in this area have well defined spatial receptive fields that are similar in their structure to receptive fields of lower visual areas. Neurons are tuned in both linear and spiral motion space and their tuning strength in both motion spaces is correlated. The position invariance of tuning for spiral motion confirmed that this tuning is not an artifact of local motion patterns being covered by the receptive field, but that the neurons really respond selectively to the global motion pattern. To understand the receptive field structure in more detail, a reverse correlation approach was used to analyze responses to a newly developed random motion stimulus (the "RC stimulus"). However, this approach was only partially able to predict neuronal responses to other stimuli. A potential reason why this method was not more successful could be that the unstructured

stimulus did not elicit strong responses from neurons that are tuned for global motion patterns. Therefore, chapter 6 introduced a second new stimulus (the "hex-stimulus") that also consisted of separate segments. In contrast to the stimulus from chapter 5, however, these segments could be independently controlled, which allows to simulate motion patterns similar to the kind of stimuli traditionally used to drive MST neurons.

Finally, chapter 7 provided results in support of the hypothesis that motion perception does not rely on a single processing pipeline, but that different areas, such as V1 and MT, vary in their importance and contribution to perception, depending on stimulus features.

How do these results fit into our current understanding of motion perception and vision more broadly? A recurrent theme throughout all the chapters of this dissertation has been the hierarchical structure of the primate visual system and the transition "from sensation to cognition". One can think of the primate visual cortex as having, roughly speaking, three stages: the first, "early" stage is essentially the primary cortex (V1), whose neurons respond to changes in low-level physical attributes of stimuli, such as luminance, contrast, and spatial frequency. The second, "mid-level" stage consists of areas V4 and MT in the ventral and dorsal pathway respectively. Neurons in these areas represent what I call the "building blocks" of perception. In the ventral pathway, for example, color constancy (rather than wavelength selectivity) is more prominent in V4 neurons than in earlier areas (Zeki, 1983). In the dorsal pathway, MT neurons respond to pattern motion (rather than component motion) similar to what observers report in psychophysical experiments (Movshon et al., 1985). In other words, what is represented by neurons in these two areas corresponds well to the colloquial terms that are used to describe components of visual images. The third, "high-level" stage consists of IT cortex in the ventral and area MST in the dorsal pathway. Neurons in these areas are concerned with global percepts, such as recognizing the faces of people around us (IT), or determining the direction we are currently moving in (MST).

Over the last seven decades, researchers have amassed a very good (though certainly not complete) understanding of the early processing of stimulus features in V1. The transition from V1 to the middle stage (V4 and MT) has seen some progress, with relatively simple, hierarchical models being able to account for responses in these areas (e.g., Simoncelli & Heeger, 1998; Rust et al., 2006; Nishimoto & Gallant, 2011). In what terms we should think about the high-level areas, however, is less clear. Chapter 5 in particular has attempted to describe MST with the same tools that have been used for low and mid-level areas, such a tuning curves and regression analyses. While

some light is shed on the response properties of MST neurons, it is clear that our results do not form a complete picture. The finding of correlated tuning strength for linear and spiral tuning in chapter 5 and strong responses to motion patterns outside the linear and spiral space reported in chapter 6, strongly suggest that neurons in this area are not easily described by a simple model with few parameters.

Perry and Fallah (2014) pointed out that the ventral pathway is often described as creating an integrated representation of perceptual objects. The dorsal pathway's role, on the other hand, has classically been described as representing individual motion features, such as direction and speed. These authors review evidence for an alternative point of view in which areas along the dorsal pathway also integrate individual features into intermediate object representations that might also incorporate information from the ventral visual pathway (Perry & Fallah, 2014). This idea, that the whole (i.e., the representation in high-level areas like MST) is greater than the sum of its parts (i.e., the representation of individual features in low- and mid-level areas), fits well with our results. For example, our findings of position invariance (chapter 5), which means that a neuron responds selectively to a global motion pattern ("the whole"), not to a specific arrangement of local motion patterns ("the sum of its parts"), is consistent with such a framework. Furthermore, the limited success of our regression analyses, which quite literally try to describe a neuron's response to a motion pattern as the sum of its responses to parts of the motion pattern, also agrees with this theory.

The big question that remains is how representations of individual features in low- and mid-level areas are integrated into the larger perceptual units that appear to be represented by high-level areas like MST. The approaches that were introduced in chapter 6 might prove to be fruitful to address this question. The hex-stimulus allows to create stimuli that are perceptually similar to integrated, "whole" stimuli (even though, technically, it is made up of individual parts). The adaptive sampling approach allows for such integrated stimuli to be flexibly constructed, based on the neuron's responses. Together, these two tools offer great potential to explore the representation of complex motion in MST.

It is important to note that this three-stage framework is a simplification that offers guidance for thinking about the visual system, but has its limitations. First, ascribing simple concepts such as "motion processing" or "color perception" to areas like MT or V4 ignores the heterogeneity of neurons in these areas. In V4, for example, estimates of the proportion of color-selective cells range from 20% (Schein et al., 1982) to 100% (Zeki, 1973). In MT, the proportion of direction selective neurons that have been reported range from

76% (Zeki, 1980) to 100% (Albright, 1984). The experiments presented in chapters 5 and 6 of this dissertation also excluded a significant number of cells from the analyses because they did not show any selective responses to motion (see Table 5.1). In general, many acute single electrode recordings in nonhuman primates suffer from a selection bias, in that often only well isolated neurons showing desirable properties (such as direction selectivity in MT) are investigated further. Because of this bias, the true proportion of neurons that show "typical" response properties may well be overestimated for many areas. Zeki (1983) himself notes: "No one should, however, read into this evidence the supposition that all the cells of the V4 complex are colour-coded [sic] as opposed to being wavelength-selective or that all cells of V1 are wavelength-selective." The role of neurons in these areas that do not share the typical features is not yet understood. The existence of non-motion sensitive neurons in MT and non-color coding neurons in V4 clearly shows that ascribing such one-dimensional roles to these areas is an oversimplification.

The second caveat for the three-stage model is that these stages do not form a one way street where an image is passed on from early to mid-level to high-level visual areas culminating in a percept of the world. Instead, there are extensive feedback connections from higher to lower areas. The *Reverse Hierarchy Theory* (RHT, Hochstein & Ahissar, 2002) proposes that visual perception starts with "vision at a glance", where what is first perceived is the output of the last stage of the visual hierarchy. "Vision with scrutiny" relies on traveling back along the visual hierarchy to reach more specialized cells with smaller receptive fields (Hochstein & Ahissar, 2002). For example, when seeing a house, we can immediately recognize it as a house ("vision at a glance"). But to perceive fine-grained details, such as an ornament on the front door ("vision with scrutiny"), activity in lower visual areas needs to be reactivated (Ahissar et al., 2009). This theory is supported by findings that for low contrast, barely visible stimuli to be perceived, they must elicit activity in frontal areas that then "ignites" a sustained pattern of activity (van Vugt et al., 2018). Our results in chapter 7 fit well into this scheme, as they indicate that the visual system is able to flexibly recruit visual areas based on stimulus features and task demands. How exactly the brain knows whether to rely on activity from one area versus another is an open question.

8.2 Limitations and future directions

All the different analysis approaches used in this dissertation – spatial receptive fields, tuning curves, reverse correlation, adaptive sampling – attempt to

describe a neuron's firing rate as a function of external stimuli. In reality, however, the only cells in the visual pathway that have direct access to the external world are the photoreceptors of the retina. All subsequent neurons, from retinal ganglion cells through the LGN, V1, and MT, all the way to MST only receive the neural activity of the preceding stage as their input. The further up one moves in the visual hierarchy, the further the recorded activity is removed from the actual stimulus. This distance between neurons in high-level visual areas and the external world could be one reason why the attempt to describe MST responses with a linear-nonlinear-Poisson (LNP) model in chapter 5 was not more successful. A number of recent papers have addressed this issue through *subunit* models in which a neuron's receptive field is divided into several subunits that can be thought of as distinct input channels to the neuron (e.g., smaller receptive fields of neurons that provide input to the recorded neuron). For retinal ganglion cells Liu et al. (2017) developed a method called spike-triggered non-negative matrix factorization that decomposes high-dimensional data, such as the spike-triggered stimulus ensemble, into a set of modules with corresponding weights. This method was able to retrieve the receptive fields of simultaneously recorded bipolar cells from the spike trains of retinal ganglion cells alone. A different approach with a similar outcome, also in the retina, uses cascaded linear-nonlinear (LN-LN) models where the stimulus is processed by multiple filters. The output of each filter is fed into a nonlinearity (first LN stage) and the outputs of the nonlinearities are linearly summed and put through another nonlinearity (second LN stage). This method led to a 53% improvement in predicting the resposnes of RGCs as compared to a one-layer LN model (Maheswaranathan et al., 2018). In V1 such a LN-LN cascade model similarly outperformed a one-layer LN model, an energy model (Adelson & Bergen, 1985), and a spike-triggered covariance (STC) model (Rust et al., 2005) in terms of efficiency and accuracy (Vintch et al., 2015). The MT model by Simoncelli and Heeger (1998) that was described in the introduction similarly describes MT as receiving input from V1-like units, rather than from the stimulus. Finally, and most relevant to the work presented here, such a cascade model has also been successfully applied to MST (Mineault et al., 2012). From a purely statistical point of view, any improvement in a model's ability to predict neural responses can be considered a success in itself. From a biological perspective, however, such models raise the question of how the subunits can be interpreted. For RGCs, it is reasonable to think of the subunits as the receptive fields of bipolar cells that provide the input to RGCs. In fact, simultaneous recordings have confirmed an overlap of subunits and bipolar cell receptive fields (2017). For a high-level visual area like MST, however, the interpretation is less clear-cut. Mineault et al. (2012) designed their

hierarchical model in a way so that the subunits resembled MT neurons in terms of tuning for motion direction and speed. However, the authors admit that they "cannot say that the subunits recovered by our model correspond exactly to the anatomical inputs received by each MST neuron". Of course MT neurons do not have direct access to the stimulus either, but would need to be characterized based on a cascade of models (e.g, receiving input from complex cells in V1 (Simoncelli & Heeger, 1998), which receive input from simple cells (Vintch et al., 2015), and so on). Ideally one would then describe neurons in high-level visual areas with five or six-layered models. However, such models would have many parameters that would require unrealistic amounts of data to it adequately.

A promising development along these lines in the ventral visual pathway has been inspired by the advances in computer vision that have been made possible by deep convolutional neural networks (CNNs; e.g., Krizhevsky, Sutskever, & Hinton, 2012). For example, the DiCarlo lab investigated how such networks, which are optimized for performance based on goal-driven learning, compared to neural activity. These models are not designed to simulate the neural pathways of the visual system but only optimized for performance in an ethologically valid task, such as image recognition. Thus, the problem of collecting adequate neural data to which the parameters could be fitted is avoided. Nevertheless, Yamins et al. (2014) found that such a model's mid- and late-stage layers were highly predictive of responses in primate visual areas V4 and IT (see Yamins & DiCarlo, 2016, for a review). These authors conclude that such performance-optimized models might offer a valuable tool in understanding neural processing (also see Kriegeskorte, 2015).

To my knowledge no comparable efforts have been made for models of the dorsal visual pathway. While the motion processing system poses some additional problems, such as constantly changing input, the computer vision community should be highly interested as motion is an important issue in one of the most exciting applications of computer vision, self-driving cars.

Another limitation of the experiments presented in this dissertation is the lack of ethological validity. As mentioned in the discussion of chapter 5, our "Reverse correlation stimulus" evoked much weaker responses than optic flow dot patterns. Given the results about multisensory integration reviewed in chapters 2 and 3, one can speculate that even the responses to optic flow dot patterns might still not be representative of "real life" MST activity in response to actual self-motion. On the other hand, natural stimuli are much more difficult to parameterize and might modulate neural activity in ways that are difficult to control for. The advantages and disadvantages of using

synthetic, highly-controlled versus natural, ethologically valid stimuli in visual neuroscience have been discussed at great length (see, for example, Rust & Movshon, 2005 and Olshausen & Field, 2005 for two opposing views). One of the major challenges in deciphering the neural code of MST will be to reconcile the opposing ends of this debate, which will, at least partially, rely on new technology (see last section). Closed-loop, adaptive sampling methods, like the one outlined in chapter 6, will be very useful in combining the complexity of natural stimuli with the statistical requirements of advanced analysis methods.

8.3 Concluding remarks

It is often recommended that scientific writing follows a "broad-narrow-broad" structure (e.g., Mensh & Kording, 2017) that starts by outlining the overarching, "big picture" problem to which the writing attempts to contribute; then describes the actual contribution in detail; and ends by embedding the findings into the big picture and providing an outlook on the open questions that are yet to be addressed in the field. I would like to end this dissertation with the "biggest" picture perspective and some ideas where the field might go in the coming years. I have alluded to these ideas in section 2.9, but would like to use this space to expand upon them in a slightly bolder and more speculative manner.

For many (though certainly not all) neuroscientists, the ultimate goal is to have a complete understanding of the human brain, both to unravel its remarkable computational power, and to have a foundation on which cures for neurological and psychiatric diseases can be built. The "perfect" experiment to achieve this would require a number of ingredients, the combination of which is, at the current point in time, unfeasible:

1. we would like to record the activity of thousands of individual neurons simultaneously during natural behavior;

2. we would like to have full information about genetically, anatomically, and physiologically defined cell types;

3. we would like to be able to manipulate neural activity at fine temporal and spatial resolution (ideally at the single cell level);

4. we would like to have perfect, quantitative measurements of all behavioral outcomes, such as limb movements and (especially in visual neuroscience) direction of gaze;

5. we would like to do all this non-invasively in human subjects.

The last point is easiest to address as non-invasive single cell recordings are simply not possible. We will therefore rely on animal models for many years to come. And while much is to be learned from simpler model organisms, there is no way around nonhuman primates when it comes to complex cognitive tasks and comparability to the human brain.

As for the other four points, great progress has been made in each of them individually: high-density microelectrode arrays already make it possible to record hundreds of neurons simultaneously (e.g., M. M. Churchland et al., 2010; Dann et al., 2016; Michaels et al., 2020) across several areas of the primate cortex. Unfortunately, such electrode arrays can only be used on the cortical surface, but not in areas like MT and MST that are located inside a sulcus, or subcortical structure. Linear electrode arrays can be inserted into the brain's sulci, but have much fewer channels. However, recently developed high-density NeuroPixel probes (Jun et al., 2017) allow many more neurons to be recorded simultaneously across brain structures in rodents (Steinmetz et al., 2019) and are beginning to be used in nonhuman primates as well (Trautmann et al., 2019; Shenoy & Kao, 2021).

It has long been known that all neurons are not created equal. The most obvious differentiation in the mammalian cortex is between glutamatergic pyramidal neurons and GABAergic interneurons, but even this is a gross oversimplification (Zeng & Sanes, 2017). Genetically encoded indicators of specific molecules, such as calcium (Ca^{2+}) (T.-W. Chen et al., 2013; Dana et al., 2019) or neurotransmitters (Feng et al., 2019; Jing et al., 2018; Marvin et al., 2013) can help to target genetically defined cell types and explore their specific roles.

Traditional interventional approaches in neuroscience include microstimulation and pharmacological inactivation. While microstimulation has a very high temporal precision, it affects all neurons around the tip of the stimulating electrode equally, with no regard for cell type or connectivity patterns among neurons. Neuropharmacology can, by definition, be used to target specific pharmacological reactions (e.g., specific neurotransmitter systems), but often works on the timescale of at least several seconds, if not minutes or hours (depending on the exact techniques being used), which is far removed from the millisecond precision with which neural circuits operate. Optogenetics combines the advantages of both methods, as it allows the excitation or inhibition of genetically defined cell populations at a millisecond timescale (Boyden et al., 2005; Deisseroth, 2011, 2015). As with NeuroPixel probes, this technique has been successfully applied by numerous labs in rodents, but has taken more time to be adapted for use with nonhuman primates (A. Afraz et al.,

2015; Acker et al., 2016; Fortuna et al., 2020; Tremblay et al., 2020).

Even simultaneous recordings and stimulations of thousands of genetically defined cells will not be able to fully explain the brain without taking behavior into consideration (Krakauer et al., 2017). However, to add behavior to the equation (and I mean this both metaphorically and literally by, for example, including measures of behavior in regression analyses) it needs to be quantified. Ideally this should be done in an automated, unbiased manner. Powerful computer vision algorithms have successfully been employed in order to track research animals' position and pose (e.g., DeepLabCut by A. Mathis et al., 2018 and MoSeq by Wiltschko et al., 2015, see M. W. Mathis & Mathis, 2020, for a review).

I want to highlight two recent publications that have combined several of these new approaches to bring us closer to the "perfect experiment" outlined above. Berger et al. (2020) combined wireless recordings from chronically implanted microelectrode arrays in three areas of the primate motor cortex with automated pose estimation in unrestrained monkeys to investigate the planning and execution of walk-and-reach movements. Even more relevant to the topics investigated in this dissertation is the work presented in a recent preprint by Mao and colleagues (2021): to explore spatial navigation, they combined chronically implanted tetrodes and single electrodes, maker-based tracking of head position, and wireless eye tracking to measure the activity of neurons in the hippocampal formation during free foraging in an open arena. As described in chapter 2, self-motion perception and spatial navigation are intimately connected and the possibility to record neural activity and track eye movements in freely moving animals will open new possibilities for investigating an area like MST.

Considering that many of the tools I have described in this section, such as optogenetics, NeuroPixels, and DeepLabCut, have only been developed in the last 15 years, there can be no doubt that these are exciting times in neuroscience.

References

Acker, L., Pino, E. N., Boyden, E. S., & Desimone, R. (2016). FEF inactivation with improved optogenetic methods. *Proceedings of the National Academy of Sciences*, *113*(46), E7297–E7306. doi: 10.1073/pnas.1610784113

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284. doi: 10.1364/JOSAA.2.000284

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525. doi: 10.1038/300523a0

Adrian, E. D. (1928). *The Basis of Sensation*. W. W. Norton & Co.

Afraz, A., Boyden, E. S., & DiCarlo, J. J. (2015). Optogenetic and pharmacological suppression of spatial clusters of face neurons reveal their causal role in face gender discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(21), 6730–6735. doi: 10.1073/pnas.1423328112

Afraz, S.-R., Kiani, R., & Esteky, H. (2006). Microstimulation of inferotemporal cortex influences face categorization. *Nature*, *442*(7103), 692–695. doi: 10.1038/nature04982

Ahissar, M., & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, *8*(10), 457–464. doi: 10.1016/j.tics.2004.08.011

Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1515), 285–299. doi: 10.1098/rstb.2008.0253

Alberts, B. B. G. T., Selen, L. P. J., & Medendorp, W. P. (2019). Age-related re-weighting of visual and vestibular cues for vertical perception. *Journal of Neurophysiology*, *121*, 1279–1288. doi: 10.1152/jn.00481.2018

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, *52*(6), 1106–1130. doi: 10.1152/jn.1984.52.6.1106

Albright, T. D., Desimone, R., & Gross, C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology*, *51*(1), 16–31. doi: 10.1152/jn.1984.51.1.16

Allen, W. E., Kauvar, I. V., Chen, M. Z., Richman, E. B., Yang, S. J., Chan, K., ... Deisseroth, K. (2017). Global Representations of Goal-Directed Behavior in Distinct Cell Types of Mouse Neocortex. *Neuron*, *94*, 891–907.e6. doi: 10.1016/j.neuron.2017.04.017

Allman, J. M., & Kaas, J. H. (1971). A representation of the visual field in the caudal third of the middle temporal gyrus of the owl monkey (Aotus trivirgatus). *Brain Research*, *31*(1), 85–105. doi: 10.1016/0006-8993(71)90635-4

Amaral, D., Insausti, R., & Cowan, W. (1983). Evidence for a direct projection from the superior temporal gyrus to the entorhinal cortex in the monkey. *Brain Research*, *275*(2), 263–277. doi: 10.1016/0006-8993(83)90987-3

Andersen, R. A., Asanuma, C., Essick, G., & Siegel, R. M. (1990). Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *The Journal of Comparative Neurology*, *296*(1), 65–113. doi: 10.1002/cne.902960106

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology*, *19*(4), 452–458. doi: 10.1016/j.conb.2009.06.008

Arnason, G. (2020). The Emergence and Development of Animal Research Ethics: A Review with a Focus on Nonhuman Primates. *Science and Engineering Ethics*, *26*(4), 2277–2293. doi: 10.1007/s11948-020-00219-z

Attwell, D., & Laughlin, S. B. (2001). An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow and Metabolism*, *21*, 1133–45. doi: 10.1097/00004647-200110000-00001

Bair, W., & Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Computation*, *8*(6), 1185–1202. doi: 10.1162/neco.1996.8.6.1185

Bakhtiari, S., Awada, A., & Pack, C. C. (2020). Influence of stimulus complexity on the specificity of visual perceptual learning. *Journal of Vision*, *20*(6), 13-13. doi: 10.1167/jov.20.6.13

Baloni, S. (2012). *Spatial, feature and temporal attentional mechanisms in visual motion processing* (Doctoral dissertation, Georg-August-Universität Göttingen). Retrieved from `http://hdl.handle.net/11858\/00-1735-0000-0001-BA0F-B`

Bao, P., She, L., Mcgill, M., & Tsao, D. Y. (2020). A map of object space in primate inferotemporal cortex. *Nature*, *583*, 103–108. doi: 10.1038/s41586-020-2350-5

Barbas, H., & Mesulam, M.-M. (1981). Organization of afferent input to subdivisions of area 8 in the rhesus monkey. *The Journal of Comparative Neurology*, *200*(3), 407–431. doi: 10.1002/cne.902000309

Barlow, H. B., & Hill, R. M. (1963). Selective Sensitivity to Direction of Movement in Ganglion Cells of the Rabbit Retina. *Science*, *139*(3553), 412–412. doi: 10.1126/science.139.3553.412

Barton, J. J., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, *58*(1), 71–78. doi: 10.1212/wnl.58.1.71

Bear, M. F., Conners, B. W., & Paradiso, M. A. (2007). *Neuroscience: Exploring the brain.* Lippincott, Williams & Wilkins.

Beardsley, S. A., & Vaina, L. M. (1998). Computational modelling of optic flow selectivity in MSTd neurons. *Network: Computation in Neural Systems*, *9*(4), 467–493. doi: 10.1088/0954-898X_9_4_005

Beauchamp, M. S., Cox, R. W., & DeYoe, E. A. (1997). Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *Journal of Neurophysiology*, *78*(1), 516–20. doi: 10.1152/jn.1997.78.1.516

Beck, J. M., Latham, P. E., & Pouget, A. (2011). Marginalization in neural circuits with divisive normalization. *The Journal of Neuroscience*, *31*(43), 15310–15319. doi: 10.1523/JNEUROSCI.1706-11.2011

Ben Hamed, S., Page, W. K., Duffy, C. J., & Pouget, A. (2003). MSTd neuronal basis functions for the population encoding of heading direction. *Journal of*

Neurophysiology, *90*(2), 549–58. doi: 10.1152/jn.00639.2002

Benda, J., Gollisch, T., Machens, C. K., & Herz, A. V. (2007). From response to stimulus: adaptive sampling in sensory physiology. *Current Opinion in Neurobiology*, *17*(4), 430–436. doi: 10.1016/j.conb.2007.07.009

Berens, P. (2009). CircStat: A MATLAB Toolbox for Circular Statistics. *Journal of Statistical Software*, *31*(10). doi: 10.18637/jss.v031.i10

Berger, M., Agha, N. S., & Gail, A. (2020). Wireless recording from unrestrained monkeys reveals motor goal encoding beyond immediate reach in frontoparietal cortex. *eLife*, *9*, e51322. doi: https://doi.org/10.7554/eLife.51322

Berger, M., Calapai, A., Stephan, V., Niessing, M., Burchardt, L., Gail, A., & Treue, S. (2018). Standardized automated training of rhesus monkeys for neuroscience research in their housing environment. *Journal of Neurophysiology*, *119*, 796–807. doi: 10.1152/jn.00614.2017

Betts, L. R., Sekuler, A. B., & Bennett, P. J. (2009). Spatial characteristics of center-surround antagonism in younger and older adults. *Journal of Vision*, *9*(1), 25–25. doi: 10.1167/9.1.25

Betts, L. R., Taylor, C. P., Sekuler, A. B., & Bennett, P. J. (2005). Aging Reduces Center-Surround Antagonism in Visual Motion Processing. *Neuron*, *45*(3), 361–366. doi: 10.1016/j.neuron.2004.12.041

Bisley, J. W., & Pasternak, T. (2000). The Multiple Roles of Visual Cortical Areas MT/MST in Remembering the Direction of Visual Motion. *Cerebral Cortex*, *10*(11), 1053–1065. doi: 10.1093/cercor/10.11.1053

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, *203*, 237–260. doi: 10.1113/jphysiol.1969.sp008862

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, *9*(2), 46–52. doi: 10.1016/j.tics.2004.12.006

Borghuis, B. G., Perge, J. A., Vajda, I., van Wezel, R. J., van de Grind, W. A., & Lankheet, M. J. (2003). The motion reverse correlation (MRC) method: A linear systems approach in the motion domain. *Journal of Neuroscience Methods*, *123*(2), 153–166. doi: 10.1016/S0165-0270(02)00347-3

Born, R. T. (2000). Center-Surround Interactions in the Middle Temporal Visual Area of the Owl Monkey. *Journal of Neurophysiology*, *84*(5), 2658–2669. doi: 10.1152/jn.2000.84.5.2658

Born, R. T., & Bradley, D. C. (2005). Structure and Function of Visual Area MT. *Annual Review of Neuroscience*, *28*(1), 157–189. doi: 10.1146/annurev.neuro.26.041002.131052

Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1992). Subcortical connections of visual areas MST and FST in macaques. *Visual Neuroscience*, *9*(3-4), 291–302. doi: 10.1017/s0952523800010701

Boussaoud, D., Ungerleider, L. G., & Desimone, R. (1990). Pathways for motion analysis: Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *The Journal of Comparative Neurology*, *296*(3), 462–495. doi: 10.1002/cne.902960311

Bowen, R. W., Pola, J., & Matin, L. (1974). Visual persistence: Effects of flash luminance, duration and energy. *Vision Research*, *14*(4), 295–303. doi: 10.1016/0042-6989(74)90079-0

Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., & Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, *8*(9), 1263–1268. doi: 10.1038/nn1525

Bradley, D. C., Maxwell, M., Andersen, R. A., Banks, M. S., & Shenoy, K. V. (1996). Mechanisms of Heading Perception in Primate Visual Cortex. *Science*, *273*, 1544–1547. doi: 10.1126/science.273.5281.1544

Bremmer, F., Churan, J., & Lappe, M. (2017). Heading representations in primates are compressed by saccades. *Nature Communications*, *8*(1), 920. doi: 10.1038/s41467-017-01021-5

Bremmer, F., Ilg, U. J., Thiele, A., Distler, C., & Hoffmann, K.-P. (1997). Eye Position Effects in Monkey Cortex . I . Visual and Pursuit-Related Activity in Extrastriate Areas MT and MST. *Journal of Neurophysiology*, *77*, 944–961. doi: 10.1152/jn.1997.77.2.944

Bremmer, F., Klam, F., Duhamel, J.-R., Ben Hamed, S., & Graf, W. (2002). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience*, *16*(8), 1569–1586. doi: 10.1046/j.1460-9568.2002.02206.x

Bremmer, F., Kubischik, M., Pekel, M., Hoffmann, K. P., & Lappe, M. (2010). Visual selectivity for heading in monkey area MST. *Experimental Brain Research*, *200*(1), 51–60. doi: 10.1007/s00221-009-1990-3

Bremmer, F., Pouget, A., & Hoffmann, K.-P. (1998). Eye position encoding in the macaque posterior parietal cortex. *European Journal of Neuroscience*, *10*(1), 153–160. doi: 10.1046/j.1460-9568.1998.00010.x

Brette, R. (2015). Philosophy of the Spike: Rate-Based vs. Spike-Based Theories of the Brain. *Frontiers in Systems Neuroscience*, *9*, 1–14. doi: 10.3389/fnsys.2015.00151

Brette, R. (2019). Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences*, *42*, e215. doi: 10.1017/S0140525X19000049

Britten, K. H. (1998). Clustering of response selectivity in the medial superior temporal area of extrastriate cortex in the macaque monkey. *Visual Neuroscience*, *15*(3), 553–558. doi: 10.1017/s0952523898153166

Britten, K. H. (2008). Mechanisms of self-motion perception. *Annual Review of Neuroscience*, *31*, 389–410. doi: 10.1146/annurev.neuro.29.051605.112953

Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responese of neurons in macaque MT. *Visual Neuroscience*, *13*, 87–100.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience*, *12*(12), 4745–4765. doi: 10.1523/jneurosci.12-12-04745.1992

Britten, K. H., & van Wezel, R. J. A. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nature Neuroscience*, *1*(1), 59–63. doi: 10.1038/259

Brochier, T., Zehl, L., Hao, Y., Duret, M., Sprenger, J., Denker, M., ... Riehle, A. (2018). Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task. *Scientific Data*, *5*(1), 180055. doi: 10.1038/sdata.2018.55

Brody, B. A. (2017). Defending animal research: An international perspective. In S. J. Armstrong & R. G. Botzler (Eds.), *The animal ethics reader* (chap. 38). Taylor & Francis.

Buffalo, E. A., Movshon, J. A., & Wurtz, R. H. (2019). From basic brain research to treating human brain disorders. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(52), 26167–26172. doi: 10.1073/pnas.1919895116

Burr, D. C., Concetta Morrone, M., & Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans.

Vision Research, *38*(12), 1731–1743.

Burr, D. C., & Thompson, P. (2011). Motion psychophysics: 1985-2010. *Vision Research*, *51*(13), 1431–56. doi: 10.1016/j.visres.2011.02.008

Buschman, T. J., & Miller, E. K. (2007). Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science*, *315*(5820), 1860–1862. doi: 10.1126/science.1138071

Busigny, T., Van Belle, G., Jemel, B., Hosein, A., Joubert, S., & Rossion, B. (2014). Face-specific impairment in holistic perception following focal lesion of the right anterior temporal lobe. *Neuropsychologia*, *56*(1), 312–333. doi: 10.1016/j.neuropsychologia.2014.01.018

Busse, L., Katzner, S., Tillmann, C., & Treue, S. (2008). Effects of attention on perceptual direction tuning curves in the human visual system. *Journal of Vision*, *8*(9), 2–2. doi: 10.1167/8.9.2

Butts, D. A., Weng, C., Jin, J., Yeh, C. I., Lesica, N. A., Alonso, J. M., & Stanley, G. B. (2007). Temporal precision in the neural code and the timescales of natural vision. , *449*(7158), 92–95. doi: 10.1038/nature06105

Buzsáki, G., & Draguhn, A. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(5679), 1926–1929. doi: 10.1126/science.1099745

Cai, D., Deangelis, G. C., & Freeman, R. D. (1997). Spatiotemporal Receptive Field Organization in the Lateral Geniculate Nucleus of Cats and Kittens. *Journal of Neurophysiology*, *78*(2), 1045–1061. doi: 10.1152/jn.1997.78.2.1045

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., & Gail, A. (2016). A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. *Behavior Research Methods*, *49*(1), 35–45. doi: 10.3758/s13428-016-0707-3

Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annual Review of Neuroscience*, *21*, 47–74. doi: 10.1146/annurev.neuro.21.1.47

Callaway, E. M. (2005). Structure and function of parallel pathways in the primate early visual system. *The Journal of Physiology*, *566*(1), 13–19. doi: 10.1113/jphysiol.2005.088047

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*, *197*(3), 551–66. doi: 10.1113/jphysiol.1968.sp008574

Carandini, M., Demb, J. B., Mante, V., Tolhurst, D. J., Dan, Y., Olshausen, B. A., ... Rust, N. C. (2005). Do We Know What the Early Visual System Does? *The Journal of Neuroscience*, *25*(46), 10577–10597. doi: 10.1523/jneurosci.3726-05.2005

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*, 51–62. doi: 10.1038/nrn3136

Casagrande, V. A. (1994). A third parallel visual pathway to primate area V1. *Trends in Neurosciences*, *17*(7), 305–310. doi: 10.1016/0166-2236(94)90065-5

Castelo-Branco, M., Formisano, E., Backes, W., Zanella, F., Neuenschwander, S., Singer, W., & Goebel, R. (2002). Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(21), 13914–13919. doi: 10.1073/pnas.202049999

Celebrini, S., & Newsome, W. (1994). Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. *The Journal of Neuroscience*, *14*(7), 4109–4124. doi: 10.1523/jneurosci.14-07

-04109.1994

Celebrini, S., & Newsome, W. T. (1995). Microstimulation of extrastriate area MST influences performance on a direction discrimination task. *Journal of Neurophysiology*, *73*(2), 437–448. doi: 10.1152/jn.1995.73.2.437

Chalk, M., Herrero, J. L., Gieselmann, M. A., Delicato, L. S., Gotthardt, S., & Thiele, A. (2010). Attention Reduces Stimulus-Driven Gamma Frequency Oscillations and Spike Field Coherence in V1. *Neuron*, *66*(1), 114–125. doi: 10.1016/j.neuron.2010.03.013

Chalmers, D. J. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, *2*(3), 200–219.

Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011a). A Comparison of Vestibular Spatiotemporal Tuning in Macaque Parietoinsular Vestibular Cortex, Ventral Intraparietal Area, and Medial Superior Temporal Area. *The Journal of Neuroscience*, *31*(8), 3082–3094. doi: 10.1523/jneurosci.4476-10.2011

Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011b). Convergence of Vestibular and Visual Self-Motion Signals in an Area of the Posterior Sylvian Fissure. *The Journal of Neuroscience*, *31*(32), 11617–11627. doi: 10.1523/jneurosci.1266-11.2011

Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011c). Representation of Vestibular and Visual Cues to Self-Motion in Ventral Intraparietal Cortex. *The Journal of Neuroscience*, *31*(33), 12036–12052. doi: 10.1523/jneurosci.0395-11.2011

Chen, A., Gu, Y., Takahashi, K., Angelaki, D. E., & DeAngelis, G. C. (2008). Clustering of Self-Motion Selectivity and Visual Response Properties in Macaque Area MSTd. *Journal of Neurophysiology*, *100*(5), 2669–2683. doi: 10.1152/jn.90705.2008

Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., ... Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300. doi: 10.1038/nature12354

Chichilnisky, E. J. (2001). A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, *12*(2), 199–213. doi: 10.1080/713663221

Chowdhury, S. A., Takahashi, K., DeAngelis, G. C., & Angelaki, D. E. (2009). Does the Middle Temporal Area Carry Vestibular Signals Related to Self-Motion? *The Journal of Neuroscience*, *29*(38), 12020–12030. doi: 10.1523/jneurosci.0004-09.2009

Churchland, A. K., Huang, X., & Lisberger, S. G. (2007). Responses of Neurons in the Medial Superior Temporal Visual Area to Apparent Motion Stimuli in Macaque Monkeys. *Journal of Neurophysiology*, *97*(1), 272–282. doi: 10.1152/jn.00941.2005

Churchland, M. M., & Lisberger, S. G. (2001). Shifts in the Population Response in the Middle Temporal Visual Area Parallel Perceptual and Motor Illusions Produced by Apparent Motion. *The Journal of Neuroscience*, *21*(23), 9387–9402. doi: 10.1523/jneurosci.21-23-09387.2001

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., ... Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, *13*(3), 369–378. doi: 10.1038/nn.2501

Cisek, P., & Kalaska, J. F. (2005). Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron*, *45*(5), 801–814. doi: 10.1016/j.neuron.2005.01.027

Cisek, P., & Kalaska, J. F. (2010). Neural Mechanisms for Interacting with a

World Full of Action Choices. *Annual Review of Neuroscience*, *33*, 269–298.

Cohen, M. R., & Maunsell, J. H. (2009). Attention improves performance primarily by reducing interneuronal correlations. *Nature Neuroscience*, *12*(12), 1594–1600. doi: 10.1038/nn.2439

Cohen, M. R., & Maunsell, J. H. (2011). Using Neuronal Populations to Study the Mechanisms Underlying Spatial and Feature Attention. *Neuron*, *70*(6), 1192–1204. doi: 10.1016/j.neuron.2011.04.029

Constantinidis, C., Funahashi, S., Lee, D., Murray, J. D., Qi, X.-L., Wang, M., & Arnsten, A. F. (2018). Persistent Spiking Activity Underlies Working Memory. *The Journal of Neuroscience*, *38*(32), 7020–7028. doi: 10.1523/jneurosci.2486-17.2018

Cottereau, B. R., Smith, A. T., Rima, S., Fize, D., Héjja-Brichard, Y., Renaud, L., ... Durand, J.-B. (2017). Processing of Egomotion-Consistent Optic Flow in the Rhesus Macaque Cortex. *Cerebral Cortex*, *27*, 330–343. doi: 10.1093/cercor/bhw412

Crapse, T. B., & Basso, M. A. (2015). Insights into decision making using choice probability. *Journal of Neurophysiology*, *114*(6), 3039–3049. doi: 10.1152/jn.00335.2015

Cronin, B., Stevenson, I. H., Sur, M., & Kording, K. P. (2010). Hierarchical Bayesian Modeling and Markov Chain Monte Carlo Sampling for Tuning-Curve Analysis. *Journal of Neurophysiology*, *103*, 591–602. doi: 10.1152/jn.00379.2009.

Crowell, J. A., Banks, M. S., Shenoy, K. V., & Andersen, R. A. (1998). Visual self-motion perception during head turns. *Nature Neuroscience*, *1*(8), 732–737. doi: 10.1038/3732

Cui, Y., Liu, L. D., Khawaja, F. A., Pack, C. C., & Butts, D. A. (2013). *Spiking activity in area MT of awake adult macaques in response to complex motion features. CRCNS.org*. doi: http://dx.doi.org/10.6080/K0X63JTX

Cumming, B. G., & Nienborg, H. (2016). Feedforward and feedback sources of choice probability in neural population responses. *Current Opinion in Neurobiology*, *37*, 126–132. doi: 10.1016/j.conb.2016.01.009

Curran, W., Beattie, L., Bilello, D., Coulter, L. A., Currie, J. A., & Leon, J. M. (2019). The direction after-effect is a global motion phenomenon. *Royal Society Open Science*, *6*(3), 190114. doi: 10.1098/rsos.190114

Dan, Y., Alonso, J.-M., Usrey, W. M., & Reid, R. C. (1998). Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nature Neuroscience*, *1*(6), 501–507. doi: 10.1038/2217

Dana, H., Sun, Y., Mohar, B., Hulse, B. K., Kerlin, A. M., Hasseman, J. P., ... Kim, D. S. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nature Methods*, *16*(7), 649–657. doi: 10.1038/s41592-019-0435-6

Dann, B., Michaels, J. A., Schaffelhofer, S., & Scherberger, H. (2016). Uniting functional network topology and oscillations in the fronto-parietal single unit network of behaving primates. *eLife*, *5*, 1–27. doi: 10.7554/eLife.15719

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. MIT Press.

De Boer, E., & Kuyper, P. (1968). Triggered Correlation. *IEEE Transactions on Biomedical Engineering*, *BME-15*(3), 169–179. doi: 10.1109/TBME.1968.4502561

DeAngelis, G. C., & Newsome, W. T. (1999). Organization of Disparity-Selective Neurons in Macaque Area MT. *The Journal of Neuroscience*, *19*(4), 1398–1415. doi: 10.1523/jneurosci.19-04

-01398.1999

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993a). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *Journal of Neurophysiology*, *69*(4), 1091–1117. doi: 10.1152/jn.1993.69.4.1091

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993b). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *Journal of Neurophysiology*, *69*(4), 1118–1135. doi: 10.1152/jn.1993.69.4.1118

DeAngelis, G. C., & Uka, T. (2003). Coding of Horizontal Disparity and Velocity by MT Neurons in the Alert Macaque. *Journal of Neurophysiology*, *89*(2), 1094–1111. doi: 10.1152/jn.00717.2002

DeCharms, R. C., & Zador, A. (2000). Neural Representation and the Cortical Code. *Annual Review of Neuroscience*, *23*(1), 613–647. doi: 10.1146/annurev.neuro.23.1.613

Deisseroth, K. (2011). Optogenetics. *Nature Methods*, *8*(1), 26–29. doi: 10.1038/nmeth.f.324

Deisseroth, K. (2015). Optogenetics: 10 years of microbial opsins in neuroscience. *Nature Neuroscience*, *18*(9), 1213–1225. doi: 10.1038/nn.4091

Dennett, D. C. (1991). *Consciousness explained*. Little, Brown & Co.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective Properties of Inferior Temporal Neurons in the Macaque. *The Journal of Neuroscience*, *4*(8), 2051–2062. doi: 10.1523/jneurosci.04-08-02051.1984

Desimone, R., & Ungerleider, L. G. (1986). Multiple visual areas in the caudal superior temporal sulcus of the macaque. *The Journal of Comparative Neurology*, *248*(2), 164–189. doi: 10.1002/cne.902480203

DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences*, *11*(8), 333–341. doi: 10.1016/j.tics.2007.06.010

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, *73*(3), 415–34. doi: 10.1016/j.neuron.2012.01.010

Diggelmann, R., Fiscella, M., Hierlemann, A., & Franke, F. (2018). Automatic spike sorting for high-density microelectrode arrays. *Journal of Neurophysiology*, *120*(6), 3155–3171. doi: 10.1152/jn.00803.2017

Dodd, J. V., Krug, K., Cumming, B. G., & Parker, A. J. (2001). Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *The Journal of Neuroscience*, *21*(13), 4809–21. doi: 10.1523/jneurosci.21-13-04809.2001

Dokka, K., DeAngelis, G. C., & Angelaki, D. E. (2015). Multisensory Integration of Visual and Vestibular Signals Improves Heading Discrimination in the Presence of a Moving Object. *The Journal of Neuroscience*, *35*(40), 13599–13607. doi: 10.1523/jneurosci.2267-15.2015

Douglas, R. M., Neve, A., Quittenbaum, J. P., Alam, N. M., & Prusky, G. T. (2006). Perception of visual motion coherence by rats and mice. *Vision Research*, *46*(18), 2842–2847. doi: 10.1016/j.visres.2006.02.025

Dubner, R., & Zeki, S. (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. doi: 10.1016/0006-8993(71)90494-X

Duffy, C. J. (1998). MST Neurons Respond to Optic Flow and Translational Movement. *Journal of Neurophysiology*, *80*(4), 1816–1827. doi:

10.1152/jn.1998.80.4.1816

Duffy, C. J., & Wurtz, R. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *The Journal of Neuroscience*, *15*(7), 5192–5208. doi: 10.1523/jneurosci.15-07-05192.1995

Duffy, C. J., & Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*(6), 1329–45. doi: 10.1152/jn.1991.65.6.1329

Duffy, C. J., & Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology*, *65*(6), 1346–1359. doi: 10.1152/jn.1991.65.6.1346

Duffy, C. J., & Wurtz, R. H. (1997). Medial Superior Temporal Area Neurons Respond to Speed Patterns in Optic Flow. *The Journal of Neuroscience*, *17*(8), 2839–2851. doi: 10.1523/jneurosci.17-08-02839.1997

Dukelow, S. P., Desouza, J. F., Culham, J. C., Van Den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology*, *86*(4), 1991–2000. doi: 10.1152/jn.2001.86.4.1991

Dürsteler, M. R., & Wurtz, R. H. (1988). Pursuit and optokinetic deficits following chemical lesions of cortical areas MT and MST. *Journal of Neurophysiology*, *60*(3), 940–965. doi: 10.1152/jn.1988.60.3.940

Eifuku, S., & Wurtz, R. H. (1998). Response to Motion in Extrastriate Area MSTl: Center-Surround Interactions. *Journal of Neurophysiology*, *80*(1), 282–296. doi: 10.1152/jn.1998.80.1.282

Elliott, T. R. (1904). On the action of adrenalin. *Journal of Physiology (London)*, *31*, 20–21. doi: 10.1113/jphysiol.1904.sp001055

Esghaei, M., Daliri, M. R., & Treue, S. (2015). Attention Decreases Phase-Amplitude Coupling, Enhancing Stimulus Discriminability in Cortical Area MT. *Frontiers in Neural Circuits*, *9*, 82. doi: 10.3389/fncir.2015.00082

Esghaei, M., Daliri, M. R., & Treue, S. (2018). Attention decouples action potentials from the phase of local field potentials in macaque visual cortical area MT. *BMC Biology*, *16*(86). doi: 10.1186/s12915-018-0551-2

Felleman, D. J., & Kaas, J. H. (1984). Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *Journal of Neurophysiology*, *52*(3), 488–513. doi: 10.1152/jn.1984.52.3.488

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47. doi: 10.1093/cercor/1.1.1-a

Feng, J., Zhang, C., Lischinsky, J. E., Jing, M., Zhou, J., Wang, H., ... Li, Y. (2019). A Genetically Encoded Fluorescent Sensor for Rapid and Specific In Vivo Detection of Norepinephrine. *Neuron*, *102*(4), 745–761.e8. doi: 10.1016/j.neuron.2019.02.037

Ferrera, V. P., Rudolph, K. K., & Maunsell, J. H. (1994). Responses of neurons in the parietal and temporal visual pathways during a motion task. *The Journal of Neuroscience*, *14*(10), 6171–6186. doi: 10.1523/jneurosci.14-10-06171.1994

Ferster, D., Chung, S., & Wheat, H. (1996). Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature*, *380*(6571), 249–252. doi: 10.1038/380249a0

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2010). Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory. *European Journal of Neuroscience*, *31*(10), 1721–1729. doi: 10.1111/j.1460-9568.2010.07207.x

Fetsch, C. R., DeAngelis, G. C., & Angelaki, D. E. (2013). Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nature Reviews Neuroscience*, *14*(6), 429–442. doi: 10.1038/nrn3503

Fetsch, C. R., Pouget, A., Deangelis, G. C., & Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience*, *15*(1), 146–154. doi: 10.1038/nn.2983

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., & Angelaki, D. E. (2009). Dynamic Reweighting of Visual and Vestibular Cues during Self-Motion Perception. *The Journal of Neuroscience*, *29*(49), 15601–15612. doi: 10.1523/jneurosci.2574-09.2009

Fetsch, C. R., Wang, S., Gu, Y., Deangelis, G. C., & Angelaki, D. E. (2007). Spatial reference frames of visual, vestibular, and multimodal heading signals in area MSTd. *The Journal of Neuroscience*, *27*(3), 700–712. doi: 10.1523/jneurosci.3553-06.2007

Fortuna, M. G., Hüer, J., Guo, H., Gruber, J., Gruber-Dujardin, E., Staiger, J. F., ... Gail, A. (2020). Histological assessment of optogenetic tools to study fronto-visual and fronto-parietal cortical networks in the rhesus macaque. *Scientific Reports*, *10*(1), 1–17. doi: 10.1038/s41598-020-67752-6

Fredericksen, R. E., Bex, P. J., & Verstraten, F. A. J. (1998). How big is a Gabor patch, and why should we care? *Journal of the Optical Society of America A*, *15*(7), 1959. doi: 10.1364/JOSAA.15.001959

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*(9), 1195–1201. doi: 10.1038/nn.2889

Freiwald, W. A., & Tsao, D. Y. (2010). Functional Compartmentalization and Viewpoint Generalization Within the Macaque Face-Processing System. *Science*, *330*, 845–852. doi: 10.1126/science.1194908

Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science*, *291*(5508), 1560–3. doi: 10.1126/science.1055465

Froehler, M. T., & Duffy, C. J. (2002). Cortical Neurons Encoding Path and Place: Where You Go Is Where You Are. *Science*, *295*(5564), 2462–2465. doi: 10.1126/science.1067426

Frost, A., & Niemeier, M. (2015). Suppression and reversal of motion perception around the time of the saccade. *Frontiers in Systems Neuroscience*, *9*(OCTOBER), 1–11. doi: 10.3389/fnsys.2015.00143

Galletti, C., Gamberini, M., Kutz, D. F., Baldinotti, I., & Fattori, P. (2005). The relationship between V6 and PO in macaque extrastriate cortex. *European Journal of Neuroscience*, *21*(4), 959–970. doi: 10.1111/j.1460-9568.2005.03911.x

Galletti, C., Gamberini, M., Kutz, D. F., Fattori, P., Luppino, G., & Matelli, M. (2001). The cortical connections of area V6: an occipito-parietal network processing visual information. *European Journal of Neuroscience*, *13*(8), 1572–1588. doi: 10.1046/j.0953-816x.2001.01538.x

Gamberini, M., Passarelli, L., Fattori, P., Zucchelli, M., Bakola, S., Luppino, G., & Galletti, C. (2009). Cortical connections of the visuomotor parietooccipital area V6Ad of the macaque monkey. *The Journal of Comparative Neurology*, *513*(6), 622–642. doi: 10.1002/cne.21980

Gandhi, N. J., & Katnani, H. A. (2011). Motor Functions of the Superior Colliculus. *Annual Review of Neuroscience*, *34*, 205–231. doi: 10.1146/annurev-neuro-061010-113728

Geesaman, B. J., & Andersen, R. A. (1996). The Analysis of Complex

Motion Patterns by Form/Cue Invariant MSTd Neurons. *The Journal of Neuroscience*, *16*(15), 4716–4732. doi: 10.1523/JNEUROSCI.16-15-04716.1996

Geesaman, B. J., Born, R. T., Andersen, R. A., & Tootell, R. B. H. (1997). Maps of complex motion selectivity in the superior temporal cortex of the alert macaque monkey: a double-label 2-deoxyglucose study. *Cerebral Cortex*, *7*(8), 749–757. doi: 10.1093/cercor/7.8.749

Gegenfurtner, K. R., Kiper, D. C., & Levitt, J. B. (1997). Functional properties of neurons in macaque area V3. *Journal of Neurophysiology*, *77*(4), 1906–23. doi: 10.1152/jn.1997.77.4.1906

Georgopoulos, A., Schwartz, A., & Kettner, R. (1986). Neuronal population coding of movement direction. *Science*, *233*(4771), 1416–1419. doi: 10.1126/science.3749885

Gibson, J. J. (1950). *The perception of the visual world.* Houghton Mifflin.

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. *Annual Review of Neuroscience*, *30*, 535–574. doi: 10.1146/annurev.neuro.29.051605.113038

Gollisch, T., & Meister, M. (2008). Rapid Neural Coding in the Retina with Relative Spike Latencies. *Science*, *319*(5866), 1108–1111. doi: 10.1126/science.1149639

Gollisch, T., & Meister, M. (2010). Eye Smarter than Scientists Believed: Neural Computations in Circuits of the Retina. *Neuron*, *65*(2), 150–164. doi: 10.1016/j.neuron.2009.12.009

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways tor perception and action. *Trends in Neurosciences*, *15*(1), 20–25. doi: 10.1016/0166-2236(92)90344-8

Graziano, M. S. A., Andersen, R. A., & Snowden, R. J. (1994). Tuning of MST neurons to spiral motions. *The Journal of Neuroscience*, *14*(1), 54–67. doi: 10.1523/JNEUROSCI.14-01-00054.1994

Grewe, B. F., Langer, D., Kasper, H., Kampa, B. M., & Helmchen, F. (2010). High-speed in vivo calcium imaging reveals neuronal network activity with near-millisecond precision. *Nature Methods*, *7*, 399–405. doi: 10.1038/nmeth.1453

Gross, C. G. (2008). Single neuron studies of inferior temporal cortex. *Neuropsychologia*, *46*(3), 841–852. doi: 10.1016/j.neuropsychologia.2007.11.009

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the Macaque. *Journal of Neurophysiology*, *35*(1), 96–111. doi: 10.1152/jn.1972.35.1.96

Grossberg, S., Mingolla, E., & Pack, C. C. (1999). A Neural Model of Motion Processing and Visual Navigation by Cortical Area MST. *Cerebral Cortex*, *9*(8), 878–895. doi: 10.1093/cercor/9.8.878

Gu, Y., Angelaki, D. E., & DeAngelis, G. C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nature Neuroscience*, *11*(10), 1201–1210. doi: 10.1038/nn.2191

Gu, Y., Chen, A., Liu, S., Fetsch, C., Yun, Y., Sunkara, A., ... Angelaki, D. (2018). *Extracellular recordings from areas MSTd and VIP of macaque monkeys during a heading discrimination task. CRCNS.org*. doi: http://dx.doi.org/10.6080/K07P8WKF

Gu, Y., DeAngelis, G. C., & Angelaki, D. E. (2007). A functional link between area MSTd and heading perception based on vestibular signals. *Nature Neuroscience*, *10*(8), 1038–1047. doi: 10.1038/nn1935

Gu, Y., DeAngelis, G. C., & Angelaki, D. E. (2012). Causal Links between Dorsal Medial Superior Temporal Area Neurons and Multisensory Heading Perception. *The Journal of Neu-*

roscience, *32*(7), 2299–2313. doi: 10.1523/jneurosci.5154-11.2012

Gu, Y., Fetsch, C. R., Adeyemo, B., DeAngelis, G. C., & Angelaki, D. E. (2010). Decoding of MSTd population activity accounts for variations in the precision of heading perception. *Neuron*, *66*(4), 596–609. doi: 10.1016/j.neuron.2010.04.026

Gu, Y., Watkins, P. V., Angelaki, D. E., & DeAngelis, G. C. (2006). Visual and Nonvisual Contributions to Three-Dimensional Heading Selectivity in the Medial Superior Temporal Area. *The Journal of Neuroscience*, *26*(1), 73–85. doi: 10.1523/jneurosci.2356-05.2006

Harth, E., & Tzanakou, E. (1974). ALOPEX: A stochastic method for determining visual receptive fields. *Vision Research*, *14*(12), 1475–1482. doi: 10.1016/0042-6989(74)90024-8

Hartline, H. (1938). The Response of Single Optic Nerve Fibers of the Vertebrate Eye to Illumination of the Retina. *American Journal of Physiology*, *121*(2), 400–415. doi: 10.1152/ajplegacy.1938.121.2.400

Harvey, M. A., Saal, H. P., Dammann, J. F., & Bensmaia, S. J. (2013). Multiplexing Stimulus Information through Rate and Temporal Codes in Primate Somatosensory Cortex. *PLoS Biology*, *11*(5), e1001558. doi: 10.1371/journal.pbio.1001558

Hassenstein, B., & Reichardt, W. (1956, oct). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers Chlorophanus. *Zeitschrift für Naturforschung B*, *11*(9-10), 513–524. doi: 10.1515/znb-1956-9-1004

Heeger, D. J. (1992). Half-squaring in responses of cat striate cells. *Visual Neuroscience*, *9*(5), 427–443. doi: 10.1017/S095252380001124X

Helmchen, F., & Denk, W. (2005). Deep tissue two-photon microscopy. *Nature Methods*, *2*, 932–940. doi: 10.1038/nmeth818

Henderickson, A. E., Wilson, J. R., & Ogren, M. P. (1978). The neuroanatomical organization of pathways between the dorsal lateral geniculate nucleus and visual cortex in old world and new world primates. *The Journal of Comparative Neurology*, *182*(1), 123–136. doi: 10.1002/cne.901820108

Herz, A. V., Meier, R., Nawrot, M. P., Schiegel, W., & Zito, T. (2008). G-Node: An integrated tool-sharing platform to support cellular and systems neurophysiology in the age of global neuroinformatics. *Neural Networks*, *21*(8), 1070–1075. doi: 10.1016/j.neunet.2008.05.011

Heuer, H. W., & Britten, K. H. (2004). Optic Flow Signals in Extrastriate Area MST: Comparison of Perceptual and Neuronal Sensitivity. *Journal of Neurophysiology*, *91*(3), 1314–1326. doi: 10.1152/jn.00637.2003

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(3), 791–804. doi: 10.1016/S0896-6273(02)01091-7

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, *117*, 500–544. doi: 10.1113/jphysiol.1952.sp004764

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, *148*, 574–591. doi: 10.1113/jphysiol.2009.174151

Hubel, D. H., & Wiesel, T. N. (1962). Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. *Journal of Physiology*, *160*, 106–154. doi: 10.1113/jphysiol.1962.sp006837

Hubel, D. H., & Wiesel, T. N. (1968). Receptive Fields and Functional Architecture of Monkey Striate Cortex. *Jour-*

nal of Physiology, *195*, 215–243. doi: 10.1113/jphysiol.1968.sp008455

Hüer, J. (2017). *Top-down attention: neural pathways in the human and non-human primate examined by electrophysiology , optogenetics and psychophysics* (Doctoral dissertation, Georg-August-Universität Göttingen). Retrieved from `http://hdl.handle.net/11858\/00-1735-0000-002E-E581-8`

Hüer, J., Fortuna, M. G., Guo, H., Schiller, L., Gail, A., Gruber, J., ... Treue, S. (2018). Optogenetic inhibition reveals a causal role of the direct anatomical connection from the FEF to extrastriate visual area MT in mediating attentional modulation in non-human primates. In *Annual Meeting of the Society for Neuroscience.* San Diego, CA, USA.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *The Journal of Neuroscience*, *22*(16), 7195–7205. doi: 10.1523/jneurosci.22-16-07195.2002

Hung, C. P., Kreiman, G., Poggio, T., & DiCarlo, J. J. (2005). Fast Readout of Object Identity from Macaque Inferior Temporal Cortex. *Science*, *310*(5749), 863–866. doi: 10.1126/science.1117593

Ilg, U. J., Schumann, S., & Thier, P. (2004). Posterior Parietal Cortex Neurons Encode Target Motion in World-Centered Coordinates. *Neuron*, *43*(1), 145–151. doi: 10.1016/j.neuron.2004.06.006

Ilg, U. J., & Thier, P. (2003). Visual Tracking Neurons in Primate Area MST Are Activated by Smooth-Pursuit Eye Movements of an Imaginary Target. *Journal of Neurophysiology*, *90*(3), 1489–1502. doi: 10.1152/jn.00272.2003

Jacobs, J., Weidemann, C. T., Miller, J. F., Solway, A., Burke, J. F., Wei, X.-x., ... Kahana, M. J. (2013). Direct recordings of grid-like neuronal activity in human spatial navigation. *Nature Neuroscience*, *16*(9), 1188–1190. doi: 10.1038/nn.3466

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, *9*(5), 690–696. doi: 10.1038/nn1691

Jing, M., Zhang, P., Wang, G., Feng, J., Mesik, L., Zeng, J., ... Li, Y. (2018). A genetically encoded fluorescent acetylcholine indicator for in vitro and in vivo studies. *Nature Biotechnology*, *36*(8). doi: 10.1038/nbt.4184

Jones, J. P., & Palmer, L. A. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, *58*(6), 1187–211. doi: 10.1152/jn.1987.58.6.1187

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., ... Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, *551*, 232–236. doi: 10.1038/nature24636

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (Eds.). (2013). *Principles of neural science* (Fifth ed.). New York: McGraw-Hill, Health Professions Division.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, *17*(11), 4302–4311. doi: 10.1523/jneurosci.17-11-04302.1997

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–2128. doi: 10.1098/rstb.2006.1934

Kaplan, E., & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences*, *83*(8), 2755–2757. doi: 10.1073/pnas.83.8.2755

Katz, B. (1969). *The release of neural transmitter substances*. Liverpool University Press.

Khawaja, F. A., Liu, L. D., & Pack, C. C. (2013). Responses of MST neurons to plaid stimuli. *Journal of Neurophysiology*, *110*(1), 63–74. doi: 10.1152/jn.00338.2012

Killian, N. J., Jutras, M. J., & Buffalo, E. A. (2012). A map of visual space in the primate entorhinal cortex. *Nature*, *491*(7426), 761–764. doi: 10.1038/nature11587

Kim, H. R., Pitkow, X., Angelaki, D. E., & DeAngelis, G. C. (2016). A simple approach to ignoring irrelevant variables by population decoding based on multisensory neurons. *Journal of Neurophysiology*, *116*(3), 1449–1467. doi: 10.1152/jn.00005.2016

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776. doi: 10.1016/j.visres.2005.10.002

Knöll, J., Pillow, J. W., & Huk, A. C. (2018). Lawful tracking of visual motion in humans, macaques, and marmosets in a naturalistic, continuous, and untrained behavioral context. *Proceedings of the National Academy of Sciences*, *115*(44), E10486–E10494. doi: 10.1073/pnas.1807192115

Koch, C. (2020). Hot or not. *Nature Human Behavior*, *4*, 991-992. doi: 10.1038/s41562-020-0925-7

Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*(I), 161–180. doi: 10.1016/0042-6989(86)90078-7

Koenderink, J. J., & van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, *22*(9), 773–791. doi: 10.1080/713819112

Kolster, H., Mandeville, J. B., Arsenault, J. T., Ekstrom, L. B., Wald, L. L., & Vanduffel, W. (2009). Visual field map clusters in macaque extrastriate visual cortex. *Journal of Neuroscience*, *29*(21), 7031–7039. doi: 10.1523/jneurosci.0518-09.2009

Kolster, H., Peeters, R., & Orban, G. A. (2010). The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. *Journal of Neuroscience*, *30*(29), 9801–9820. doi: 10.1523/jneurosci.2069-10.2010

Komatsu, H., & Wurtz, R. H. (1988a). Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *Journal of Neurophysiology*, *60*(2), 621–644. doi: 10.1152/jn.1988.60.2.621

Komatsu, H., & Wurtz, R. H. (1988b). Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *Journal of Neurophysiology*, *60*(2), 580–603. doi: 10.1152/jn.1988.60.2.580

Komatsu, H., & Wurtz, R. H. (1989). Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. *Journal of Neurophysiology*, *62*(1), 31–47. doi: 10.1152/jn.1989.62.1.31

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, *93*(3), 480–490. doi: 10.1016/j.neuron.2016.12.041

Krauzlis, R. J., Lovejoy, L. P., & Zénon, A. (2013). Superior Colliculus and Visual Spatial Attention. *Annual Review of Neuroscience*, *36*(1), 165–182. doi: 10.1146/annurev-neuro-062012-170249

Kriegeskorte, N. (2015). Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annual Review of Vision Science*, *1*(1), 417–446. doi: 10.1146/annurev-vision-082114-035447

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems* (Vol. 25). Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf

Kuffler, S. W. (1953). Discharge Patterns and Functional Organization of Mammalian Retina. *Journal of Neurophysiology*, *16*(1), 37–68. doi: 10.1152/jn.1953.16.1.37

Kusunoki, M., Moutoussis, K., & Zeki, S. (2006). Effect of Background Colors on the Tuning of Color-Selective Cells in Monkey Area V4. *Journal of Neurophysiology*, *95*(5), 3047–3059. doi: 10.1152/jn.00597.2005

Lagae, L., Maes, H., Raiguel, S., Xiao, D. K., & Orban, G. A. (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. *Journal of Neurophysiology*, *71*(5), 1597–1626. doi: 10.1152/jn.1994.71.5.1597

Lagae, L., Raiguel, S., & Orban, G. A. (1993). Speed and direction selectivity of macaque middle temporal neurons. *Journal of Neurophysiology*, *69*(1), 19–39. doi: 10.1152/jn.1993.69.1.19

Lappe, M., Bremmer, F., Pekel, M., Thiele, A., & Hoffmann, K.-P. (1996). Optic flow processing in monkey STS: a theoretical and experimental approach. *The Journal of Neuroscience*, *16*(19), 6265–85. doi: 10.1523/jneurosci.16-19-06265.1996

Lappe, M., Bremmer, F., & Van Den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, *3*(9), 329–336. doi: 10.1016/S1364-6613(99)01364-9

Lappe, M., & Rauschecker, J. (1993a). Computation of heading direction from optic flow in visual cortex. In S. Hanson, J. Cowan, & C. Giles (Eds.), *Advances in neural information processing systems* (Vol. 5).

Lappe, M., & Rauschecker, J. P. (1993b). A Neural Network for the Processing of Optic Flow from Ego-Motion in Man and Higher Mammals. *Neural Computation*, *5*(3), 374–391. doi: 10.1162/neco.1993.5.3.374

Lauwers, K., Saunders, R., Vogels, R., Vandenbussche, E., & Orban, G. A. (2000). Impairment in motion discrimination tasks is unrelated to amount of damage to superior temporal sulcus motion areas. *The Journal of Comparative Neurology*, *420*(4), 539–557. doi: 10.1002/(SICI)1096-9861(20000515)420:4⟨539::AID-CNE10⟩3.0.CO;2-3

Layton, O. W., & Fajen, B. R. (2016). A Neural Model of MST and MT Explains Perceived Object Motion during Self-Motion. *Journal of Neuroscience*, *36*(31), 8093–8102. doi: 10.1523/jneurosci.4593-15.2016

Lee, B. B. (2011). Visual pathways and psychophysical channels in the primate. *Journal of Physiology*, *589*(1), 41–47. doi: 10.1113/jphysiol.2010.192658

Lennie, P. (2003). The Cost of Cortical Computation. *Current Biology*, *13*, 493–497. doi: 10.1016/S

Lewi, J., Butera, R., & Paninski, L. (2009). Sequential Optimal Design of Neurophysiology Experiments. *Neural Computation*, *21*(3), 619–687. doi: 10.1162/neco.2008.08-07-594

Lewis, J. W., & Van Essen, D. C. (2000). Mapping of architectonic subdivisions in the macaque monkey, with emphasis on parieto-occipital cortex. *The Journal of Compar-*

ative Neurology, *428*(1), 79–111. doi: 10.1002/1096-9861(20001204)428:1⟨79::AID-CNE7⟩3.0.CO;2-Q

Lisberger, S. G. (2015). Visual Guidance of Smooth Pursuit Eye Movements. *Annual Review of Vision Science*, *1*(1), 447–468. doi: 10.1146/annurev-vision-082114-035349

Liu, J. K., Schreyer, H. M., Onken, A., Rozenblit, F., Khani, M. H., Krishnamoorthy, V., ... Gollisch, T. (2017). Inference of neuronal functional circuitry with spike-triggered non-negative matrix factorization. *Nature Communications*, *8*(1), 149. doi: 10.1038/s41467-017-00156-9

Liu, L. D., Haefner, R. M., & Pack, C. C. (2016). A neural basis for the spatial suppression of visual motion perception. *eLife*, *5*, 1–20. doi: 10.7554/eLife.16167.001

Liu, L. D., & Pack, C. C. (2017). The contribution of area MT to visual motion perception depends on training. *Neuron*, *95*(2), 436–446. doi: 10.1016/j.neuron.2017.06.024

Loewi, O. (1921). Über humorale Übertragbarkeit der Herznervenwirkung. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, *189*, 239–242. doi: 10.1007/BF01738910

Logan, D. J., & Duffy, C. J. (2006). Cortical area MSTd combines visual cues to represent 3-D self-movement. *Cerebral Cortex*, *16*(10), 1494–1507. doi: 10.1093/cercor/bhj082

Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural Mechanisms of Spatial Selective Attention in Areas V1 , V2 , and V4 of Macaque Visual Cortex. *Journal of Neurophysiology*, *77*, 24–42. doi: 10.1152/jn.1997.77.1.24

Lundqvist, M., Herman, P., & Miller, E. K. (2018). Working Memory: Delay Activity, Yes! Persistent Activity? Maybe Not. *The Journal of Neuroscience*, *38*(32), 7013–7019. doi: 10.1523/jneurosci.2485-17.2018

Luo, J., He, K., Andolina, I. M., Li, X., Yin, J., Chen, Z., ... Wang, W. (2019). Going with the flow: the neural mechanisms underlying illusions of complex-flow motion. *The Journal of Neuroscience*, *39*(14), 2664–2685. doi: 10.1523/jneurosci.2112-18.2019

Maamoun, A. (2018). *Receptive Field Characterization in MSTd Neurons* (Doctoral dissertation, Georg-August-Universität Göttingen). Retrieved from `http://hdl.handle.net/11858/\00-1735-0000-002E-E537-0`

Machens, C. K. (2002). Adaptive Sampling by Information Maximization. *Physical Review Letters*, *88*(22), 228104. doi: 10.1103/PhysRevLett.88.228104

Machens, C. K., Gollisch, T., Kolesnikova, O., & Herz, A. V. (2005). Testing the Efficiency of Sensory Coding with Optimal Stimulus Ensembles. *Neuron*, *47*(3), 447–456. doi: 10.1016/j.neuron.2005.06.015

Maheswaranathan, N., Kastner, D. B., Baccus, S. A., & Ganguli, S. (2018). Inferring hidden structure in multilayered neural circuits. *PLOS Computational Biology*, *14*(8), e1006291. doi: 10.1371/journal.pcbi.1006291

Malek, N., Treue, S., Khayat, P., & MartinezTrujillo, J. (2017). Distracter suppression dominates attentional modulation of responses to multiple stimuli inside the receptive fields of middle temporal neurons. *European Journal of Neuroscience*, *46*(12), 2844–2858. doi: 10.1111/ejn.13764

Manning, T. S., & Britten, K. H. (2019). Retinal Stabilization Reveals Limited Influence of Extraretinal Signals on Heading Tuning in the Medial Superior Temporal Area. *The Journal of Neuroscience*, *39*(41), 8064–8078. doi: 10.1523/jneurosci.0388-19.2019

Mao, D., Avila, E., Caziot, B., Laurens, J., Dickman, J. D., & Angelaki, D. E.

(2021). Spatial modulation of hippocampal activity in freely moving macaques. *Neuron*, *in press*. doi: 10.1016/j.neuron.2021.09.032

Markov, N. T., Ercsey-Ravasz, M., Van Essen, D. C., Knoblauch, K., Toroczkai, Z., & Kennedy, H. (2013). Cortical high-density counterstream architectures. *Science*, *342*(6158). doi: 10.1126/science.1238406

Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., ... Kennedy, H. (2014). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cerebral Cortex*, *24*(1), 17–36. doi: 10.1093/cercor/bhs270

Marques, T., Summers, M. T., Fioreze, G., Fridman, M., Dias, R. F., Feller, M. B., & Petreanu, L. (2018). A Role for Mouse Primary Visual Cortex in Motion Perception. *Current Biology*, *28*(11), 1703–1713. doi: 10.1016/j.cub.2018.04.012

Marr, D. (2010). *Vision*. Cambridge, Mass.: MIT Press.

Martinez-Trujillo, J. C., & Treue, S. (2002). Attentional Modulation Strength in Cortical Area MT Depends on Stimulus Contrast. *Neuron*, *35*, 365–370. doi: 10.1016/s0896-6273(02)00778-x

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Current Biology*, *14*, 744–751. doi: 10.1016/j.cub.2004.04.028

Martinez-Trujillo, J. C., Tsotsos, J. K., Simine, E., Pomplun, M., Wildes, R., Treue, S., ... Hopf, J.-M. (2005). Selectivity for speed gradients in human area MT/V5. *NeuroReport*, *16*(5), 435–438. doi: 10.1097/00001756-200504040-00004

Marvin, J. S., Borghuis, B. G., Tian, L., Cichon, J., Harnett, M. T., Akerboom, J., ... Looger, L. L. (2013). An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nature Methods*, *10*(2), 162–170. doi: 10.1038/nmeth.2333

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. doi: 10.1038/s41593-018-0209-y

Mathis, M. W., & Mathis, A. (2020). Deep learning tools for the measurement of animal behavior in neuroscience. *Current Opinion in Neurobiology*, *60*, 1–11. doi: 10.1016/j.conb.2019.10.008

Maunsell, J. H., & Cook, E. P. (2002). The role of attention in visual processing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *357*(1424), 1063–1072. doi: 10.1098/rstb.2002.1107

Maunsell, J. H., Nealey, T. A., & DePriest, D. D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience*, *10*(10), 3323–3334. doi: 10.1523/jneurosci.10-10-03323.1990

Maunsell, J. H., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, *10*, 363–401. doi: 10.1146/annurev.neuro.10.1.363

Maunsell, J. H., & Van Essen, D. C. (1983a). The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of Neuroscience*, *3*(12), 2563–2586. doi: 10.1523/jneurosci.03-12-02563.1983

Maunsell, J. H., & Van Essen, D. C. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, *49*(5), 1148–1167. doi: 10.1152/

jn.1983.49.5.1148

Maunsell, J. H., & Van Essen, D. C. (1983c). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, *49*(5), 1127–1147. doi: 10.1152/jn.1983.49.5.1127

Mauss, A. S., Vlasits, A., Borst, A., & Feller, M. (2017). Visual Circuits for Direction Selectivity. *Annual Review of Neuroscience*, *40*(1), 211–230. doi: 10.1146/annurev-neuro-072116-031335

May, P. J. (2006). The mammalian superior colliculus: Laminar structure and connections. *Progress in Brain Research*, *151*, 321–378. doi: 10.1016/S0079-6123(05)51011-2

McAdams, C. J., & Maunsell, J. H. (1999). Effects of Attention on Orientation-Tuning Functions of Single Neurons in Macaque Cortical Area V4. *The Journal of Neuroscience*, *19*(1), 431–441. doi: 10.1523/JNEUROSCI.19-01-00431.1999

Mechler, F., & Ringach, D. L. (2002). On the classification of simple and complex cells. *Vision Research*, *42*(8), 1017–1033. doi: 10.1016/S0042-6989(02)00025-1

Mehrpour, V., Martinez-Trujillo, J. C., & Treue, S. (2020). Attention amplifies neural representations of changes in sensory input at the expense of perceptual accuracy. *Nature Communications*, *11*(1), 2128. doi: 10.1038/s41467-020-15989-0

Meister, M., Pine, J., & Baylor, D. A. (1994). Multi-neuronal signals from the retina: acquisition and analysis. *Journal of Neuroscience Methods*, *51*(1), 95–106. doi: 10.1016/0165-0270(94)90030-2

Mendoza-Halliday, D., Torres, S., & Martinez-Trujillo, J. C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nature Neuroscience*, *17*(9), 1255–1262. doi: 10.1038/nn.3785

Mensh, B., & Kording, K. P. (2017). Ten simple rules for structuring papers. *PLoS Computational Biology*, *13*(9), e1005619. doi: 10.1371/journal.pcbi.1005619

Michaels, J. A., Schaffelhofer, S., Agudelo-Toro, A., & Scherberger, H. (2020). A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping. *Proceedings of the National Academy of Sciences*, *117*(50), 32124–32135. doi: 10.1073/pnas.2005087117

Micheli-Tzanakou, E. (1983). Visual receptive fields and clustering. *Behavior Research Methods & Instrumentation*, *15*(6), 553–560. doi: 10.3758/BF03203721

Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology*, *55*(6), 1328–39. doi: 10.1152/jn.1986.55.6.1328

Mikellidou, K., Frijia, F., Montanaro, D., Greco, V., Burr, D. C., & Morrone, M. C. (2018). Cortical BOLD responses to moderate- and high-speed motion in the human visual cortex. *Scientific Reports*, *8*(1), 1–12. doi: 10.1038/s41598-018-26507-0

Miller, E. K., & Cohen, J. D. (2001). An Integrative Theory of Prefrontal Cortex Function. *Annual Review of Neuroscience*, *24*(1), 167–202. doi: 10.1146/annurev.neuro.24.1.167

Miller, E. K., Erickson, C. A., & Desimone, R. (1996). Neural Mechanisms of Visual Working Memory in Prefrontal Cortex of the Macaque. *The Journal of Neuroscience*, *16*(16), 5154–5167. doi: 10.1523/jneurosci.16-16-05154.1996

Mineault, P. J., Khawaja, F. A., Butts, D. A., & Pack, C. C. (2012). Hierarchical processing of complex motion along the primate dorsal visual

pathway. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(16), E972–E980. doi: 10.1073/pnas.1115685109

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: two cortical pathways. *Trends in Neurosciences*, *6*, 414–417. doi: 10.1016/0166-2236(83)90190-X

Mitchell, A. S., Thiele, A., Petkov, C. I., Roberts, A., Robbins, T. W., & Schultz, W. (2018). Continued need for non-human primate neuroscience research. *Current Biology*, *28*(20), R1186–R1187. doi: 10.1016/j.cub.2018.09.029

Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature*, *421*, 370–373. doi: 10.1038/nature01285.

Moore, T., & Fallah, M. (2001). Control of eye movements and spatial attention. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(3), 1273–6. doi: 10.1073/pnas.021549498

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, *229*(4715), 782–784. doi: 10.1126/science.4023713

Morgan, M. L., DeAngelis, G. C., & Angelaki, D. E. (2008). Multisensory Integration in Macaque Visual Cortex Depends on Cue Reliability. *Neuron*, *59*(4), 662–673. doi: 10.1016/j.neuron.2008.06.024

Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, *376*, 507–509. doi: 10.1038/376507a0

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience*, *3*(12), 1322–1328. doi: 10.1038/81860

Moser, E. I., Kropff, E., & Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annual Review of Neuroscience*, *31*, 69–89. doi: 10.1146/annurev.neuro.31.061307.090723

Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, *70*(3), 909–919. doi: 10.1152/jn.1993.70.3.909

Movshon, J. A. (2013). Three comments on Teller's "bridge locus". *Visual neuroscience*, *30*(5-6), 219–22. doi: 10.1017/S0952523813000527

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The Analysis of Moving Visual Patterns. In C. Chagas, R. Gattass, & C. G. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117–151). New York: Springer.

Movshon, J. A., & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *The Journal of Neuroscience*, *16*(23), 7733–7741. doi: 10.1523/jneurosci.16-23-07733.1996

Müller, J. (1838). *Handbuch der physiologie des menschen für vorlesungen.*. J. Hölscher.

Nandy, A. S., Nassi, J. J., & Reynolds, J. H. (2017). Laminar Organization of Attentional Modulation in Macaque Visual Area V4. *Neuron*, *93*(1), 235–246. doi: 10.1016/j.neuron.2016.11.029

Nealey, T., & Maunsell, J. H. (1994). Magnocellular and parvocellular contributions to the responses of neurons in macaque striate cortex. *The Journal of Neuroscience*, *14*(4), 2069–2079. doi: 10.1523/JNEUROSCI.14-04-02069.1994

Nelder, J. A., & Mead, R. (1965). A Simplex

Method for Function Minimization. *The Computer Journal*, *7*(4), 308-313. doi: 10.1093/comjnl/7.4.308

Nelken, I., Prut, Y., Vaadia, E., & Abeles, M. (1994). In search of the best stimulus: An optimization procedure for finding efficient stimuli in the cat auditory cortex. *Hearing Research*, *72*(1-2), 237–253. doi: 10.1016/0378-5955(94)90222-4

Neri, P., & Levi, D. M. (2006). Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Research*, *46*(16), 2465–2474. doi: 10.1016/j.visres.2006.02.002

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, *341*(6237), 52–54. doi: 10.1038/341052a0

Newsome, W. T., & Paré, E. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *The Journal of Neuroscience*, *8*(6), 2201–2211. doi: 10.1523/jneurosci.08-06-02201.1988

Newsome, W. T., Wurtz, R., Dursteler, M., & Mikami, A. (1985). Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *The Journal of Neuroscience*, *5*(3), 825–840. doi: 10.1523/jneurosci.05-03-00825.1985

Newsome, W. T., Wurtz, R. H., & Komatsu, H. (1988). Relation ofcortical areasMTandMST to pursuit eye movements: II. Differentiation of retinal from extraretinal inputs. *Journal of Neurophysiology*, *60*(2), 604–620. doi: 10.1152/jn.1988.60.2.604

Niell, C. M., & Stryker, M. P. (2010). Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron*, *65*, 472–479. doi: 10.1016/j.neuron.2010.01.033

Niknam, K., Akbarian, A., Clark, K., Zamani, Y., Noudoost, B., & Nategh, N. (2018). *The pre-, peri-, and post-saccadic responses of MT neurons in response to random probe stimuli flashed on screen. CRCNS.org*. doi: http://dx.doi.org/10.6080/K0FB514J

Nishida, S. (2011). Advancement of motion psychophysics: Review 2001 2010. *Journal of Vision*, *11*(5), 1–53. doi: 10.1167/11.5.11.Introduction

Nishida, S., Kawabe, T., Sawayama, M., & Fukiage, T. (2018). Motion perception: From detection to interpretation. *Annual Review of Vision Science*, *4*, 501–523. doi: 10.1146/annurev-vision-091517-034328

Nishimoto, S., & Gallant, J. L. (2011). A Three-Dimensional Spatiotemporal Receptive Field Model Explains Responses of Area MT Neurons to Naturalistic Movies. *The Journal of Neuroscience*, *31*(41), 14551–14564. doi: 10.1523/jneurosci.6801-10.2011

Nishimoto, S., & Gallant, J. L. (2018). *Extracellular recordings from area MT of awake macaques in response to naturalistic movies. CRCNS.org*. doi: http://dx.doi.org/10.6080/K0DN4374

O'Connor, D. H., Fukui, M. M., Pinsk, M. a., & Kastner, S. (2002). Attention modulates responses in the human lateral geniculate nucleus. *Nature Neuroscience*, *5*(11), 1203–1209. doi: 10.1038/nn957

O'Craven, K. M., Rosen, B. R., Kwong, K. K., Treisman, A. M., & Savoy, R. L. (1997). Voluntary attention modulates fMRI activity in human MT-MST. *Neuron*, *18*(4), 591–8. doi: 10.1016/s0896-6273(00)80300-1

Ohshiro, T., Angelaki, D. E., & Deangelis, G. C. (2011). A normalization model of multisensory integration. *Nature Neuroscience*, *14*(6), 775–782. doi: 10.1038/nn.2815

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2017). A Neural Signature of Divisive Normalization at the Level of Multisensory Integration in Primate

Cortex. *Neuron*, *95*(2), 399–411.e8. doi: 10.1016/j.neuron.2017.06.043

Olshausen, B. A., & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, *14*(4), 481–487. doi: 10.1016/j.conb.2004.07.007

Olshausen, B. A., & Field, D. J. (2005). How Close Are We to Understanding V1? *Neural Computation*, *17*(8), 1665–1699. doi: 10.1162/0899766054026639

Ono, S., & Mustari, M. J. (2006). Extraretinal Signals in MSTd Neurons Related to Volitional Smooth Pursuit. *Journal of Neurophysiology*, *96*(5), 2819–2825. doi: 10.1152/jn.00538.2006

Orban, G. A. (2008). Higher Order Visual Processing in Macaque Extrastriate Cortex. *Physiological Reviews*, *88*(1), 59–89. doi: 10.1152/physrev.00008.2007

Orban, G. A., Dupont, P., De Bruyn, B., Vandenberghe, R., Rosier, A., & Mortelmans, L. (1998). Human brain activity related to speed discrimination tasks. *Experimental Brain Research*, *122*(1), 9–22. doi: 10.1007/s002210050486

Orban, G. A., Lagae, L., Raiguel, S., Xiao, D., & Maes, H. (1995). The speed tuning of medial superior temporal (MST) cell responses to optic-flow components. *Perception*, *24*, 269–285.

Orban, G. A., Lagae, L., Verri, A., Raiguel, S., Xiao, D., Maes, H., & Torre, V. (1992). First-order analysis of optical flow in monkey brain. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(7), 2595–9. doi: 10.1073/pnas.89.7.2595

Orban, G. A., Saunders, R. C., & Vandenbussche, E. (1995). Lesions of the Superior Temporal Cortical Motion Areas Impair Speed Discrimination in the Macaque Monkey. *European Journal of Neuroscience*, *7*(11), 2261–2276. doi: 10.1111/j.1460-9568.1995.tb00647.x

Orban, G. A., Van Essen, D. C., & Vanduffel, W. (2004). Comparative mapping of higher visual areas in monkeys and humans. *Trends in Cognitive Sciences*, *8*(7), 315–324. doi: 10.1016/j.tics.2004.05.009

Pack, C. C., Berezovskii, V. K., & Born, R. T. (2001). Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature*, *414*(6866), 905–908. doi: 10.1038/414905a

Pack, C. C., & Born, R. T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature*, *409*(6823), 1040–1042. doi: 10.1038/35059085

Pack, C. C., Grossberg, S., & Mingolla, E. (2001). A Neural Model of Smooth Pursuit Control and Motion Perception by Cortical Area MST. *Journal of Cognitive Neuroscience*, *13*(1), 102–120. doi: 10.1162/089892901564207

Page, W. K., & Duffy, C. J. (1999). MST neuronal responses to heading direction during pursuit eye movements. *Journal of Neurophysiology*, *81*(2), 596–610. doi: 10.1152/jn.1999.81.2.596

Page, W. K., & Duffy, C. J. (2018). Path perturbation detection tasks reduce MSTd neuronal self-movement heading responses. *Journal of Neurophysiology*, *119*(1), 124–133. doi: 10.1152/jn.00958.2016

Page, W. K., Sato, N., Froehler, M. T., Vaughn, W., & Duffy, C. J. (2015). Navigational path integration by cortical neurons: origins in higher-order direction selectivity. *Journal of Neurophysiology*, *113*(6), 1896–1906. doi: 10.1152/jn.00197.2014

Palmer, S. E. (1999). *Vision Science – Photon to Phenomenology*. Cambridge, MA: MIT Press.

Paninski, L., Pillow, J., & Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. In P. Cisek, T. Drew, &

J. F. Kalaska (Eds.), *Progress in brain research* (Vol. 165, pp. 493–507). Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S0079612306650310 doi: 10.1016/S0079-6123(06)65031-0

Paninski, L., Pillow, J. W., & Simoncelli, E. P. (2004). Maximum Likelihood Estimation of a Stochastic Integrate-and-Fire Neural Encoding Model. *Neural Computation*, *16*(12), 2533–2561. doi: 10.1162/0899766042321797

Panzeri, S., Harvey, C. D., Piasini, E., Latham, P. E., & Fellin, T. (2017). Cracking the Neural Code for Sensory Perception by Combining Statistics, Intervention, and Behavior. *Neuron*, *93*(3), 491–507. doi: 10.1016/j.neuron.2016.12.036

Paolini, M., Distler, C., Bremmer, F., Lappe, M., & Hoffmann, K.-P. (2000). Responses to continuously changing optic flow in area MST. *Journal of Neurophysiology*, *84*(2), 730–743. doi: 10.1152/jn.2000.84.2.730

Park, M., & Pillow, J. W. (2011). Receptive Field Inference with Localized Priors. *PLoS Computational Biology*, *7*(10), e1002219. doi: 10.1371/journal.pcbi.1002219

Parker, A. J., & Newsome, W. T. (1998). Sense and the Single Neuron: Probing the Physiology of Perception. *Annual Review of Neuroscience*, *21*(1), 227–277. doi: 10.1146/annurev.neuro.21.1.227

Passarelli, L., Rosa, M. G. P., Gamberini, M., Bakola, S., Burman, K. J., Fattori, P., & Galletti, C. (2011). Cortical Connections of Area V6Av in the Macaque: A Visual-Input Node to the Eye/Hand Coordination System. *Journal of Neuroscience*, *31*(5), 1790–1801. doi: 10.1523/jneurosci.4784-10.2011

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience*, *6*(2), 97–107. doi: 10.1038/nrn1603

Paxinos, G., Huang, X.-F., & Toga, A. W. (2000). *The rhesus monkey brain in stereotaxic coordinates.* San Diego: Academic Press.

Pekel, M., Lappe, M., Bremmer, F., Thiele, A., & Hoffmann, K.-P. (1996). Neuronal responses in the motion pathway of the macaque monkey to natural optic flow stimuli. *NeuroReport*, *7*, 884–888. doi: 10.1097/00001756-199603220-00010

Percie du Sert, N., Hurst, V., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., ... Würbel, H. (2020). The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLOS Biology*, *18*(7), e3000410. doi: 10.1371/journal.pbio.3000410

Perge, J. a., Borghuis, B. G., Bours, R. J. E., Lankheet, M. J. M., & van Wezel, R. J. A. (2005). Dynamics of directional selectivity in MT receptive field centre and surround. *European Journal of Neuroscience*, *22*(8), 2049–2058. doi: 10.1111/j.1460-9568.2005.04363.x

Perrone, J. A., & Thiele, A. (2001). Speed skills: measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience*, *4*(5), 526–532. doi: 10.1038/87480

Perrone, J. A., & Thiele, A. (2002). A model of speed tuning in MT neurons. *Vision Research*, *42*(8), 1035–1051. doi: 10.1016/S0042-6989(02)00029-9

Perry, C. J., & Fallah, M. (2014). Feature integration and object representations along the dorsal stream visual hierarchy. *Frontiers in Computational Neuroscience*, *8*(August), 1–17. doi: 10.3389/fncom.2014.00084

Peuskens, H., Sunaert, S., Dupont, P., Van Hecke, P., & Orban, G. A. (2001). Human brain regions involved in heading estimation. *The Journal of Neuroscience*, *21*(7), 2451–2461. doi: 10.1523/jneurosci.21-07-02451.2001

Pfefferle, D., Plümer, S., Burchardt, L., Treue, S., & Gail, A. (2018). Assessment of stress responses in rhesus macaques (Macaca mulatta) to daily routine procedures in system neuroscience based on salivary cortisol concentrations. *PLos One*, *13*(1), e0190190. doi: https://doi.org/10.1371/journal.pone.0190190

Pillow, J. W. (2007). Likelihood-Based Approaches to Modeling the Neural Code. In K. Doya, S. Ishii, A. Pouget, & R. P. N. Rao (Eds.), *Bayesian brain: Probabilistic approaches to neural coding* (pp. 53–70). MIT Press. doi: 10.7551/mitpress/9780262042383.003.0003

Pillow, J. W., Paninski, L., Uzzell, V. J., Simoncelli, E. P., & Chichilnisky, E.-J. (2005). Prediction and Decoding of Retinal Ganglion Cell Responses with a Probabilistic Spiking Model. *The Journal of Neuroscience*, *25*(47), 11003–11013. doi: 10.1523/jneurosci.3305-05.2005

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J., & Simoncelli, E. P. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, *454*(7207), 995–999. doi: 10.1038/nature07140

Pillow, J. W., & Simoncelli, E. P. (2006). Dimensionality reduction in neural models: An information-theoretic generalization of spike-triggered average and covariance analysis. *Journal of Vision*, *6*(4), 9. doi: 10.1167/6.4.9

Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (2002). *Numerical Recipes in C: The Art of Scientific Computing* (2nd ed.). Cambridge University Press.

Priebe, N. J. (2016). Mechanisms of Orientation Selectivity in the Primary Visual Cortex. *Annual Review of Vision Science*, *2*(1), 85–107. doi: 10.1146/annurev-vision-111815-114456

Priebe, N. J., Cassanello, C. R., & Lisberger, S. G. (2003). The neural representation of speed in macaque area MT/V5. *The Journal of Neuroscience*, *23*(13), 5650–5661. doi: 10.1523/JNEUROSCI.23-13-05650.2003

Priebe, N. J., & Ferster, D. (2012). Mechanisms of Neuronal Computation in Mammalian Visual Cortex. *Neuron*, *75*(2), 194–208. doi: 10.1016/j.neuron.2012.06.011

Raiguel, S., Hulle, M. M., Xiao, D.-K., Marcar, V. L., & Orban, G. A. (1995). Shape and Spatial Distribution of Receptive Fields and Antagonistic Motion Surrounds in the Middle Temporal Area (V5) of the Macaque. *European Journal of Neuroscience*, *7*(10), 2064–2082. doi: 10.1111/j.1460-9568.1995.tb00629.x

Raiguel, S., Van Hulle, M. M., Xiao, D.-K., Marcar, V. L., Lagae, L., & Orban, G. A. (1997). Size and shape of receptive fields in the medial superior temporal area (MST) of the macaque. *NeuroReport*, *8*(12), 2803–2808. doi: 10.1097/00001756-199708180-00030

Regan, D., & Beverley, K. I. (1978). Looming detectors in the human visual pathway. *Vision Research*, *18*(4), 415–421. doi: 10.1016/0042-6989(78)90051-2

Regan, T. (1983). *The Case for Animal Rights*. University of California Press.

Reid, R. C., & Alonso, J.-M. (1995). Specificity of monosynaptic connections from thalamus to visual cortex. *Nature*, *378*(6554), 281–284. doi: 10.1038/378281a0

Reid, R. C., & Shapley, R. M. (1992). Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. *Nature*, *356*(6371), 716–718. doi: 10.1038/356716a0

Richert, M., Albright, T. D., & Krekelberg, B. (2013). The Complex Structure of Receptive Fields in the Middle Temporal Area. *Frontiers in Systems Neuroscience*, *7*(March), 2. doi: 10.3389/fnsys.2013.00002

Rieke, F., Warland, D., De Ruyter Van Steveninck, R. R., & Bialek, W. (1999). *Spikes: Exploring the Neural Code*. Cambridge, MA: MIT Press. doi: 10.1016/S0065-230X(09)04001-9

Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, *2*(11), 1019–1025. doi: 10.1038/14819

Ringach, D. L. (1998). Tuning of orientation detectors in human vision. *Vision Research*, *38*(7), 963–972. doi: 10.1016/S0042-6989(97)00322-2

Ringach, D. L., & Shapley, R. M. (2004). Reverse correlation in neurophysiology. *Cognitive Science*, *28*(2), 147–166. doi: 10.1016/j.cogsci.2003.11.003

Robson, J. G. (1966). Spatial and Temporal Contrast-Sensitivity Functions of the Visual System. *Journal of the Optical Society of America*, *56*(8), 1141. doi: 10.1364/JOSA.56.001141

Rodieck, R. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, *5*(12), 583–601. doi: 10.1016/0042-6989(65)90033-7

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, *395*(6700), 376–381. doi: 10.1038/26475

Roelfsema, P. R., & Treue, S. (2014). Basic Neuroscience Research with Nonhuman Primates: A Small but Indispensable Component of Biomedical Research. *Neuron*, *82*(6), 1200–1204. doi: 10.1016/j.neuron.2014.06.003

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of Neurophysiology*, *30*(4), 769–793. doi: 10.1152/jn.1967.30.4.769

Roy, J.-P., Komatsu, H., & Wurtz, R. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *The Journal of Neuroscience*, *12*(7), 2478–2492. doi: 10.1523/jneurosci.12-07-02478.1992

Roy, J.-P., & Wurtz, R. H. (1990). The role of disparity-sensitive cortical neurons in signalling the direction of self-motion. *Nature*, *348*(6297), 160–162. doi: 10.1038/348160a0

Royden, C. S., Banks, M. S., & Crowell, J. A. (1992). The perception of heading during eye movements. *Nature*, *360*(6404), 583–585. doi: 10.1038/360583a0

Royden, C. S., & Hildreth, E. C. (1999). Differential effects of shared attention on perception of heading and 3-D object motion. *Perception & Psychophysics*, *61*(1), 120–133. doi: 10.3758/BF03211953

Rozzi, S., Calzavara, R., Belmalih, A., Borra, E., Gregoriou, G. G., Matelli, M., & Luppino, G. (2006). Cortical Connections of the Inferior Parietal Cortical Convexity of the Macaque Monkey. *Cerebral Cortex*, *16*(10), 1389–1417. doi: 10.1093/cercor/bhj076

Rudolph, K., & Pasternak, T. (1999). Transient and permanent deficits in motion perception after lesions of cortical areas MT and MST in the macaque monkey. *Cerebral Cortex*, *9*(1), 90–100. doi: 10.1111/j.1460-9568.1995.tb00629.x

Ruff, D. A., & Cohen, M. R. (2014). Attention can either increase or decrease spike count correlations in visual cortex. *Nature Neuroscience*, *17*(11), 1591–1598. doi: 10.1038/nn.3835

Russell, W. M. S., & Burch, R. L. (1959). *The principles of humane experimental technique*. Methuen.

Rust, N. C., & DiCarlo, J. J. (2010). Selectivity and Tolerance ("Invariance") Both Increase as Visual Information Propagates from Cortical Area V4 to IT. *The Journal of Neuroscience*, *30*(39), 12978–12995. doi: 10.1523/jneurosci.0179-10.2010

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience*, *9*(11), 1421–1431. doi: 10.1038/nn1786

Rust, N. C., & Movshon, J. A. (2005). In praise of artifice. *Nature Neuroscience*, *8*(12), 1647–1650. doi: 10.1038/nn1606

Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal Elements of Macaque V1 Receptive Fields. *Neuron*, *46*(6), 945–956. doi: 10.1016/j.neuron.2005.05.021

Sack, A. T., Kohler, A., Linden, D. E., Goebel, R., & Muckli, L. (2006). The temporal characteristics of motion processing in hMT/V5+: Combining fMRI and neuronavigated TMS. *NeuroImage*, *29*(4), 1326–1335. doi: 10.1016/j.neuroimage.2005.08.027

Saito, H.-A., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *The Journal of Neuroscience*, *6*(1), 145–157. doi: 10.1523/JNEUROSCI.06-01-00145.1986

Saleem, A. B., Ayaz, A. I., Jeffery, K. J., Harris, K. D., & Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. *Nature Neuroscience*, *16*, 1864–1869. doi: 10.1038/nn.3567

Salzman, C. D., Britten, K. H., & Newsome, W. T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, *346*(6280), 174–177. doi: 10.1038/346174a0

Salzman, C. D., Murasugi, C., Britten, K., & Newsome, W. (1992). Microstimulation in visual area MT: effects on direction discrimination performance. *The Journal of Neuroscience*, *12*(6), 2331–2355. doi: 10.1523/jneurosci.12-06-02331.1992

Sasaki, R., Angelaki, D. E., & DeAngelis, G. C. (2017). Dissociation of self-motion and object motion by linear population decoding that approximates marginalization. *The Journal of Neuroscience*, *37*(46), 11204–11219. doi: 10.1523/JNEUROSCI.1177-17.2017

Sasaki, R., Angelaki, D. E., & DeAngelis, G. C. (2019). Processing of object motion and self-motion in the lateral subdivision of the medial superior temporal area in macaques. *Journal of Neurophysiology*, *121*(4), 1207–1221. doi: 10.1152/jn.00497.2018

Sauseng, P., Conci, M., Wild, B., & Geyer, T. (2015). Predictive coding in visual search as revealed by cross-frequency EEG phase synchronization. *Frontiers in Psychology*, *6*, 1655. doi: 10.3389/fpsyg.2015.01655

Schein, S. J., Marrocco, R. T., & de Monasterio, F. M. (1982). Is there a high concentration of color-selective cells in area V4 of monkey visual cortex? *Journal of Neurophysiology*, *47*(2), 193–213. doi: 10.1152/jn.1982.47.2.193

Schwartz, O., Pillow, J. W., Rust, N. C., & Simoncelli, E. P. (2006). Spike-triggered neural characterization. *Journal of Vision*, *6*(4), 13. doi: 10.1167/6.4.13

Schwiedrzik, C. M., & Freiwald, W. A. (2017). High-Level Prediction Signals in a Low-Level Area of the Macaque Face-Processing Hierarchy. *Neuron*, *96*(1), 89–97.e5. doi: 10.1016/j.neuron.2017.09.007

Searle, J. (2004). *Mind. A Brief Introduction.* Oxford University Press.

Serrano-Pedraza, I., Hogg, E. L., & Read, J. C. A. (2011). Spatial nonhomogeneity of the antagonistic surround in motion perception. *Journal of Vision*, *11*(2), 3–3. doi: 10.1167/11.2.3

Shadlen, M. N. (2006). Rate versus Temporal Coding Models. In L. Nadel (Ed.), *Encyclopedia of cognitive sci-*

ence (pp. 819–825). Chichester: John Wiley & Sons, Ltd. doi: 10.1002/0470018860.s00372

Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, *4*(4), 569–579. doi: 10.1016/0959-4388(94)90059-0

Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: seeing and deciding. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 628–633. doi: 10.1073/pnas.93.2.628

Sharpee, T. O. (2013). Computational Identification of Receptive Fields. *Annual Review of Neuroscience*, *36*(1), 103–120. doi: 10.1146/annurev-neuro-062012-170253

Shenoy, K. V., & Kao, J. C. (2021). Measurement, manipulation and modeling of brain-wide neural population dynamics. *Nature Communications*, *12*(1), 633. doi: 10.1038/s41467-020-20371-1

Sherrington, C. S. (1906). *The Integrative Action of the Nervous System*. New Haven: Yale University Press.

Shiori, S., & Cavanagh, P. (1989). Saccadic suppression of low-level motion. *Vision Research*, *29*(8), 915–928. doi: 10.1016/0042-6989(89)90106-5

Sillito, A. M. (1975). The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat. *The Journal of Physiology*, *250*(2), 305–329. doi: 10.1113/jphysiol.1975.sp011056

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743–761. doi: 10.1016/S0042-6989(97)00183-1

Singer, P. (1975). *Animal Liberation*. Harper Collins.

Skottun, B. C., De Valois, R. L., Grosof, D. H., Movshon, J. A., Albrecht, D. G., & Bonds, A. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision Research*, *31*(7-8), 1078–1086. doi: 10.1016/0042-6989(91)90033-2

Smith, A. T., Wall, M. B., Williams, A. L., & Singh, K. D. (2006). Sensitivity to optic flow in human cortical areas MT and MST. *European Journal of Neuroscience*, *23*(2), 561–569. doi: 10.1111/j.1460-9568.2005.04526.x

Smith, M. (2020). *V4 Utah Array Plaid Movie Data. Figshare*. doi: https://doi.org/10.6084/m9.figshare.12269513.v1

Snowden, R. J., & Milne, A. B. (1996). The effects of adapting to complex motions: Position invariance and tuning to spiral motions. *Journal of Cognitive Neuroscience*, *8*(5), 435–452. doi: 10.1162/jocn.1996.8.5.435

Snowden, R. J., Treue, S., & Andersen, R. A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research*, *88*(2), 389–400. doi: 10.1007/BF02259114

Softky, W., & Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *The Journal of Neuroscience*, *13*(1), 334–350. doi: 10.1523/jneurosci.13-01-00334.1993

Solomon, S. G., Tailby, C., Cheong, S. K., & Camp, A. J. (2010). Linear and Nonlinear Contributions to the Visual Sensitivity of Neurons in Primate Lateral Geniculate Nucleus. *Journal of Neurophysiology*, *104*(4), 1884–1898. doi: 10.1152/jn.01118.2009

Squire, R. F., Noudoost, B., Schafer, R. J., & Moore, T. (2013). Prefrontal contributions to visual selective attention. *Annual Review of Neuroscience*, *36*, 451–66. doi: 10.1146/annurev-neuro-062111-150439

Steinmetz, N. A., Zatka-Haas, P., Carandini, M., & Harris, K. D. (2019). Distributed coding of choice, action, and engagement across the mouse

brain. *Nature*, *576*, 266–273. doi: 10.1038/s41586-019-1787-x

Stoewer, A., Kellner, C., Benda, J., Wachtler, T., & Grewe, J. (2014). File format and library for neuroscience data and metadata. In *Neuroinformatics 2014*. Leiden, Netherlands. doi: 10.3389/conf.fninf.2014.18.00027

Strong, S. L., Silson, E. H., Gouws, A. D., Morland, A. B., & McKeefry, D. J. (2017). A direct demonstration of functional differences between subdivisions of human V5/MT+. *Cerebral Cortex*, *27*(1), 1–10. doi: 10.1093/cercor/bhw362

Tadin, D., & Lappin, J. S. (2005). Optimal size for perceiving motion decreases with contrast. *Vision Research*, *45*(16), 2059–2064. doi: 10.1016/j.visres.2005.01.029

Tadin, D., Lappin, J. S., Gilroy, L. A., & Blake, R. (2003). Perceptual consequences of centre surround antagonism in visual motion processing. *Nature*, *424*, 312–315. doi: 10.1038/nature01800

Tadin, D., Park, W. J., Dieter, K. C., Melnick, M. D., Lappin, J. S., & Blake, R. (2019). Spatial suppression promotes rapid figure-ground segmentation of moving objects. *Nature Communications*, *10*(1), 2732. doi: 10.1038/s41467-019-10653-8

Tadin, D., Silvanto, J., Pascual-Leone, A., & Battelli, L. (2011). Improved motion perception and impaired spatial suppression following disruption of cortical area MT/V5. *Journal of Neuroscience*, *31*(4), 1279–1283. doi: 10.1523/jneurosci.4121-10.2011

Takahashi, K., Gu, Y., May, P. J., Newlands, S. D., DeAngelis, G. C., & Angelaki, D. E. (2007). Multimodal Coding of Three-Dimensional Rotation and Translation in Area MSTd: Comparison of Visual and Vestibular Selectivity. *The Journal of Neuroscience*, *27*(36), 9742–9756. doi: 10.1523/jneurosci.0817-07.2007

Tanaka, K. (1983). Cross-correlation analysis of geniculostriate neuronal relationships in cats. *Journal of Neurophysiology*, *49*(6), 1303–1318. doi: 10.1152/jn.1983.49.6.1303

Tanaka, K., Hikosaka, K., Saito, H.-A., Yukie, M., Fukada, Y., & Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *The Journal of Neuroscience*, *6*(1), 134–144. doi: 10.1523/jneurosci.06-01-00134.1986

Tanaka, K., & Saito, H.-A. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*(3), 626–641. doi: 10.1152/jn.1989.62.3.626

Tanaka, K., Sugita, Y., Moriya, M., & Saito, H.-A. (1993). Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. *Journal of Neurophysiology*, *69*(1), 128–142. doi: 10.1152/jn.1993.69.1.128

Taylor, W., & Vaney, D. I. (2003). New directions in retinal research. *Trends in Neurosciences*, *26*(7), 379–385. doi: 10.1016/S0166-2236(03)00167-X

Teller, D. Y. (1984). Linking propositions. *Vision Research*, *24*(10), 1233–1246. doi: 10.1016/0042-6989(84)90178-0

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*(6582), 520–522. doi: 10.1038/381520a0

Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., ... Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *The Journal of Neuroscience*, *15*(4), 3215–3230. doi: 10.1523/jneurosci.15-04-03215.1995

Touryan, J., Lau, B., & Dan, Y. (2002). Isola-

tion of Relevant Visual Features from Random Stimuli for Cortical Complex Cells. *The Journal of Neuroscience*, *22*(24), 10811–10818. doi: 10.1523/jneurosci.22-24-10811.2002

Trautmann, E. M., Stavisky, S. D., Lahiri, S., Ames, K. C., Kaufman, M. T., O'Shea, D. J., ... Shenoy, K. V. (2019). Accurate Estimation of Neural Population Dynamics without Spike Sorting. *Neuron*, *103*(2), 292–308.e4. doi: 10.1016/j.neuron.2019.05.003

Tremblay, S., Acker, L., Afraz, A., Albaugh, D. L., Amita, H., Andrei, A. R., ... Platt, M. L. (2020). An Open Resource for Non-human Primate Optogenetics. *Neuron*, 1–16. doi: 10.1016/j.neuron.2020.09.027

Treue, S. (2001). Neural correlates of attention in primate visual cortex. *Trends in Neurosciences*, *24*(5), 295–300. doi: 10.1016/S0166-2236(00)01814-2

Treue, S. (2003). Climbing the cortical ladder from sensation to perception. *Trends in Cognitive Sciences*, *7*(11), 469–471. doi: 10.1016/j.tics.2003.09.003

Treue, S., & Andersen, R. A. (1996). Neural responses to velocity gradients in macaque cortical area MT. *Visual Neuroscience*, *13*(4), 797–804. doi: 10.1017/S095252380000866X

Treue, S., & Lemon, R. (2021). The indispensable contribution of non-human primates to biomedical research. In L. Robinson & A. Weiss (Eds.), *Nonhuman primate welfare.* Cham: Springer International Publishing AG.

Treue, S., & Martinez-Trujillo, J. C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*(6736), 575–9. doi: 10.1038/21176

Treue, S., & Maunsell, J. H. (1996). Attentional modulation of visual processing in cortical areas MT and MST. *Nature*, *382*, 539–541. doi: 10.1038/382539a0

Treue, S., & Maunsell, J. H. (1999). Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *The Journal of Neuroscience*, *19*(17), 7591–602. doi: 10.1523/jneurosci.19-17-07591.1999

Treue, S., Snowden, R. J., & Andersen, R. A. (1993). The effect of transiency on perceived velocity of visual patterns: a case of temporal capture. *Vision Research*, *33*(5-6), 791–798. doi: 10.1016/0042-6989(93)90198-6

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A Cortical Region Consisting Entirely of Face-Selective Cells. *Science*, *311*(5761), 670–674. doi: 10.1126/science.1119983

Tye, K. M. (2018). Neural Circuit Motifs in Valence Processing. *Neuron*, *100*(2), 436–452. doi: 10.1016/j.neuron.2018.10.001

Tzanakou, E., Michalak, R., & Harth, E. (1979). The Alopex process: Visual receptive fields by response feedback. *Biological Cybernetics*, *35*(3), 161–174. doi: 10.1007/BF00337061

Ungerleider, L. G., & Desimone, R. (1986). Cortical connections of visual area MT in the macaque. *The Journal of Comparative Neurology*, *248*(2), 190–222. doi: 10.1002/cne.902480204

Vaina, L. M. (1998). Complex Motion Perception and its Deficits. *Current Opinion in Neurobiology*, *8*, 494–502. doi: 10.1016/S0959-4388(98)80037-8

Van Essen, D. C., Glasser, M. F., Dierker, D. L., & Harwell, J. (2012). Cortical parcellations of the macaque monkey analyzed on surface-based atlases. *Cerebral Cortex*, *22*(10), 2227–2240. doi: 10.1093/cercor/bhr290

Van Essen, D. C., Maunsell, J. H., & Bixby, J. L. (1981). The middle temporal visual area in the macaque: Myeloarchitecture, connections, functional properties and topographic organization. *The Journal of Comparative Neurol-*

ogy, *199*(3), 293–326. doi: 10.1002/cne.901990302

van Vugt, B., Dagnino, B., Vartak, D., Safaai, H., Panzeri, S., Dehaene, S., & Roelfsema, P. R. (2018). The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*, *360*(6388), 537–542. doi: 10.1126/science.aar7186

Verhoef, B.-E., & Maunsell, J. H. (2016). Attention operates uniformly throughout the classical receptive field and the surround. *eLife*, *5*. doi: 10.7554/eLife.17256

Vidyasagar, T., Pei, X., & Volgushev, M. (1996). Multiple mechanisms underlying the orientation selectivity of visual cortical neurones. *Trends in Neurosciences*, *19*(7), 272–277. doi: 10.1016/S0166-2236(96)20027-X

Vintch, B., Movshon, J. A., & Simoncelli, E. P. (2015). A Convolutional Subunit Model for Neuronal Responses in Macaque V1. *The Journal of Neuroscience*, *35*(44), 14829–14841. doi: 10.1523/jenurosci.2815-13.2015

Viswanathan, S., Jayakumar, J., & Vidyasagar, T. R. (2011). Role of feedforward geniculate inputs in the generation of orientation selectivity in the cat's primary visual cortex. *The Journal of Physiology*, *589*(9), 2349–2361. doi: 10.1113/jphysiol.2010.202317

Wald, G. (1968). The molecular basis of visual excitation. *Nature*, *219*, 800–807. doi: 10.1038/219800a0

Wall, M. B., & Smith, A. T. (2008). The Representation of Egomotion in the Human Brain. *Current Biology*, *18*(3), 191–194. doi: 10.1016/j.cub.2007.12.053

Wallach, H. (1987). Perceiving a stable environment when one moves. *Annual Review of Psychology*, *38*, 1–27. doi: 10.1146/annurev.ps.38.020187.000245

Wallisch, P., & Movshon, J. A. (2019). Responses of neurons in macaque MT to unikinetic plaids. *Journal of Neurophysiology*, *122*(5), 1937–1945. doi: 10.1152/jn.00486.2019

Wang, M., Montanède, C., Chandrasekaran, C., Peixoto, D., Shenoy, K. V., & Kalaska, J. F. (2019). Macaque dorsal premotor cortex exhibits decision-related activity only when specific stimulusresponse associations are known. *Nature Communications*, *10*. doi: 10.1038/s41467-019-09460-y

Wang, R. (1995). A simple competitive account of some response properties of visual neurons in area MSTd. *Neural Computation*, *7*(2), 290–306. doi: 10.1162/neco.1995.7.2.290

Warren, P. A., & Rushton, S. K. (2007). Perception of object trajectory: Parsing retinal motion into self and object movement components. *Journal of Vision*, *7*(11), 2–2. doi: 10.1167/7.11.2

Warren, W. H., & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature*, *336*, 162–163. doi: 10.1038/336162a0

Warren, W. H., Morris, M. W., & Kalish, M. (1988). Perception of translational heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 646–660. doi: 10.1037//0096-1523.14.4.646

Wässle, H. (2004). Parallel processing in the mammalian retina. *Nature Reviews Neuroscience*, *5*(10), 747–757. doi: 10.1038/nrn1497

Wiesel, T. N. (1959). Recording Inhibition and Excitation in the Cat's Retinal Ganglion Cells with Intracellular Electrodes. *Nature*, *183*(4656), 264–265. doi: 10.1038/183264a0

Wiesel, T. N., & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, *29*(6), 1115–1156. doi: 10.1152/jn.1966.29.6.1115

Wild, B. (2018). How Does the Brain Tell Self-Motion from Object Motion? *The Journal of Neuroscience*, *38*(16), 3875–3877. doi: 10.1523/jneurosci

.0039-18.2018

Wild, B., Maamoun, A., Mayr, Y., Brockhausen, R., & Treue, S. (2021). *Electrophysiological dataset from macaque visual cortical area MST in response to a novel motion stimulus. G-Node* `https://gin.g-node.org/cnl-dpz/macaque_MST_map1`. doi: 10.12751/g-node.d8yhh8

Wild, B., & Treue, S. (2021a). Comparing the influence of stimulus size and contrast on the perception of moving gratings and random dot patterns - A registered report protocol. *PLoS One*, *16*, e0253067. doi: 10.1371/journal.pone.0253067

Wild, B., & Treue, S. (2021b). Primate Extrastriate Cortical Area MST: A Gateway between Sensation and Cognition. *Journal of Neurophysiology*, *125*, 1851-1882. doi: 10.1152/jn.00384.2020

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. doi: 10.1038/sdata.2016.18

Williams, Z. M., Elfar, J. C., Eskandar, E. N., Toth, L. J., & Assad, J. A. (2003). Parietal activity and the perceived direction of ambiguous apparent motion. *Nature Neuroscience*, *6*(6), 616–623. doi: 10.1038/nn1055

Wilson, M. E., & Toyne, M. J. (1970). Retino-tectal and cortico-tectal projections in Macaca mulatta. *Brain Research*, *24*(3), 395–406. doi: 10.1016/0006-8993(70)90181-2

Wiltschko, A. B., Johnson, M. J., Iurilli, G., Peterson, R. E., Katon, J. M., Pashkovski, S. L., ... Datta, S. R. (2015). Mapping Sub-Second Structure in Mouse Behavior. *Neuron*, *88*(6), 1121–1135. doi: 10.1016/j.neuron.2015.11.031

Wolbers, T., Wiener, J. M., Mallot, H. A., & Buchel, C. (2007). Differential Recruitment of the Hippocampus, Medial Prefrontal Cortex, and the Human Motion Complex during Path Integration in Humans. *The Journal of Neuroscience*, *27*(35), 9408–9416. doi: 10.1523/jneurosci.2146-07.2007

Wu, M. C.-K., David, S. V., & Gallant, J. L. (2006). Complete Functional Characterization of Sensory Neurons by System Identification. *Annual Review of Neuroscience*, *29*(1), 477–505. doi: 10.1146/annurev.neuro.29.051605.113024

Xue, C., Kaping, D., Ray, S. B., Krishna, B. S., & Treue, S. (2017). Spatial Attention Reduces Burstiness in Macaque Visual Cortical Area MST. *Cerebral Cortex*, *27*(1), 83–91. doi: 10.1093/cercor/bhw326

Yamane, Y., Carlson, E. T., Bowman, K. C., Wang, Z., & Connor, C. E. (2008). A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nature Neuroscience*, *11*(11), 1352–1360. doi: 10.1038/nn.2202

Yamins, D. L. K., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience*, *19*(3), 356–365. doi: 10.1038/nn.4244

Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences*, *111*(23), 8619–8624. doi: 10.1073/pnas.1403112111

Yazdani, P., Read, J. C. A., Whittaker, R. G., & Trevelyan, A. J. (2017). Assessment of epilepsy using noninvasive visual psychophysics tests of surround suppression. *Physiological Reports*, *5*(5), 1–11. doi: 10.14814/phy2.13079

Yazdani, P., Serrano-Pedraza, I., Whittaker, R. G., Trevelyan, A., & Read, J. C. A. (2015). Two common psychophysical measures of surround suppression

reflect independent neuronal mechanisms. *Journal of Vision*, *15*(13), 21. doi: 10.1167/15.13.21

Zaidel, A., DeAngelis, G. C., & Angelaki, D. E. (2017). Decoupled choice-driven and stimulus-related activity in parietal neurons may be misrepresented by choice probabilities. *Nature Communications*, *8*(1), 715. doi: 10.1038/s41467-017-00766-3

Zehl, L., Jaillet, F., Stoewer, A., Grewe, J., Sobolev, A., Wachtler, T., ... Grün, S. (2016). Handling Metadata in a Neurophysiology Laboratory. *Frontiers in Neuroinformatics*, *10*(26). doi: 10.3389/fninf.2016.00026

Zeki, S. M. (1973). Colour coding in rhesus monkey prestriate cortex. *Brain Research*, *53*(2), 422–427. doi: 10.1016/0006-8993(73)90227-8

Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *The Journal of Physiology*, *236*(3), 549–573. doi: 10.1113/jphysiol.1974.sp010452

Zeki, S. M. (1980). The response properties of cells in the middle temporal area (area MT) of owl monkey visual cortex. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *207*(1167), 239–248. doi: 10.1098/rspb.1980.0022

Zeki, S. M. (1983). Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience*, *9*(4), 741–765. doi: 10.1016/0306-4522(83)90265-8

Zemel, R. S., & Sejnowski, T. J. (1998). A Model for Encoding Multiple Object Motions and Self-Motion in Area MST of Primate Visual Cortex. *The Journal of Neuroscience*, *18*(1), 531–547. doi: 10.1523/jneurosci.18-01-00531.1998

Zeng, H., & Sanes, J. R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. *Nature Reviews Neuroscience*, *18*(9), 530–546. doi: 10.1038/nrn.2017.85

Zhou, H., & Desimone, R. (2011). Feature-Based Attention in the Frontal Eye Field and Area V4 during Visual Search. *Neuron*, *70*(6), 1205–1217. doi: 10.1016/j.neuron.2011.04.032

Zihl, J., Von Cramon, D., & Mai, N. (1983). Selective Disturbance of Movement Vision After Bilateral Brain Damage. *Brain*, *106*, 313–340. doi: 10.1093/brain/106.2.313

Zuo, Y., Safaai, H., Notaro, G., Mazzoni, A., Panzeri, S., & Diamond, M. E. (2015). Complementary Contributions of Spike Timing and Spike Rate to Perceptual Decisions in Rat S1 and S2 Cortex. *Current Biology*, *25*(3), 357–363. doi: 10.1016/j.cub.2014.11.065

Curriculum Vitae

Education

2021	**Doctor rerum naturalium (Dr. rer. nat.)** **Göttingen University**, Germany
2015	**Master of Science (M.Sc.) in Neuro-Cognitive Psychology** **Ludwig-Maximilians-University (LMU) Munich**, Germany
2013	**Bachelor of Science (B.Sc.) in Psychology** **Heidelberg University**, Germany

Publications

- Wild, B., Treue, S. (2021a). Comparing the influence of stimulus size and contrast on the perception of moving gratings and random dot patterns - A registered report protocol. *PLoS One*, *16*, e0253067. doi: 10.1371/journal.pone.0253067

- Wild, B., Treue, S. (2021b). Primate Extrastriate Cortical Area MST: A Gateway between Sensation and Cognition. *Journal of Neurophysiology*, *125*, 1851-1882. doi: 10.1152/jn.00384.2020

- Wild, B. (2018). How Does the Brain Tell Self-Motion from Object Motion? *The Journal of Neuroscience*, *38*(16), 3875–3877. doi: 10.1523/jneurosci.0039-18.2018

- Sauseng, P., Conci, M., Wild, B., Geyer, T. (2015). Predictive coding in visual search as revealed by cross-frequency EEG phase synchronization. *Frontiers in Psychology*, *6*, 1655. doi: 10.3389/fpsyg.2015.01655