



Neural representation of depth from motion parallax in the mouse visual cortex

Yiran He

UNIVERSITY COLLEGE LONDON

Division of Biosciences

Submitted to University College London (UCL) in partial fulfilment of
the requirements for the award of the degree of Doctor of Philosophy.

Primary supervisor: **Dr. Petr Znamenskiy**

Secondary supervisor: **Prof. Neil Burgess**

Examining committee: **Prof. Isaac Bianco, Prof. Riccardo
Beltramo**

Word count: **55,059**

Thesis submission date: **29th January 2025**

Declaration

I, Yiran He, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Part of the thesis built upon findings that I have previously disseminated in a preprint on bioRxiv (He et al., 2024).

Dedication

To my beloved grandfather, a wise, kind, and honourable electrical engineer who inspired my love for science.

致我最亲爱的外公，一位智慧而慈爱的科学工作者，感谢您点亮了我心中的科学之光。

Abstract

To understand the three-dimensional structure of the world and guide behaviour, the brain must infer depth from two-dimensional retinal images. Motion parallax is an important visual cue for depth perception, relying on visual motion generated by movement of the observer. However, the neural mechanisms underlying depth perception from motion parallax remain poorly understood. Using mice as the model organism, I first established the visual cliff as a robust behavioural assay to study depth perception in mice. Using the visual cliff, I found that depth perception in mice does not depend on prior visual experience. Next, I investigated the neural mechanisms underlying depth perception from motion parallax in the mouse primary visual cortex (V1). As V1 neurons are broadly modulated by locomotion, I hypothesized that the integration of visual and locomotion-related signals in V1 neurons enables depth estimation from motion parallax. Using two-photon calcium imaging in mice navigating a three-dimensional virtual reality environment, I identified a large fraction of V1 neurons tuned to virtual depth from motion parallax. These depth-selective responses arose from conjunctive coding of visual motion and self-motion speeds. Moreover, many of these neurons responded selectively to visual stimuli presented at a specific retinotopic location and virtual depth, demonstrating that V1 neuronal responses could be characterised by three-dimensional receptive fields during active locomotion. These results demonstrate that V1 neurons create a depth map of the three-dimensional visual space by representing a wide range of virtual depths across retinotopic locations.

Impact Statement

Depth perception is crucial for animals' survival in this three-dimensional world. However, how the visual system parses the organisation of three-dimensional visual scenes to infer depth remains poorly understood. Using mice as the model organism, this thesis significantly advances our understanding of the neural mechanisms underlying depth perception.

First, using the visual cliff assay, it demonstrates that depth perception is innate in mice and does not require prior visual experience. This discovery, along with evidence of innate depth perception in many other species, suggests that the neural mechanisms underlying depth perception are likely evolutionarily conserved across animals. Therefore, this finding positions the mouse visual system and depth perception as ideal models for future studies to investigate the link between gene expression and cortical function.

Next, this thesis focuses on investigating the neural mechanisms underlying depth perception from motion parallax – one of the most important sources of depth information for rodents and many other animals. Using two-photon calcium imaging to record neuronal activity in mice navigating a three-dimensional virtual reality environment, it reveals that neurons in the primary visual cortex (V1) encode a depth map of the three-dimensional visual space during active locomotion of the animal. Traditionally, V1 was thought to represent two-dimensional visual features from retinal images. However, this study shows that V1 neurons explicitly encode three-dimensional locations of visual cues. It reshapes our understanding of how the visual system parses the structure of real-world three-dimensional environments, challenging the traditional view of V1 as a collection of two-dimensional Gabor filters.

This study also shows that the representation of depth by V1 neurons arises from by integrating visual motion and self-motion signals, which provides a functional explanation for the widespread locomotion-related modulation in V1 neurons. Since the discovery that locomotion determines the gain of responses of mouse V1 neurons, the functional role of this modulation in visual processing has been debated. This study suggests that locomotion-dependent modulation in V1 enables depth estimation from motion parallax. Furthermore, the innate nature of depth perception in mice highlights their value as a model to inform the neural mechanisms in other species.

The insights gained from this work have the potential to inform the development of biologically inspired algorithms in computer vision and robotics. By elucidating how the

brain estimates depth from motion parallax, this research can guide the development of navigation and depth estimation systems for autonomous vehicles, drones, and other artificial agents operating in complex environments.

In summary, this thesis has broad implications for both advancing basic neuroscience and inspiring interdisciplinary innovations.

Acknowledgements

Ralph Waldo Emerson once said, *"Do not go where the path may lead, go instead where there is no path and leave a trail."* My PhD journey has been a challenging yet rewarding path. There were "the best of times", sparked by the Eureka moments and the joy of discovery; there were not so good of times, marked by mysterious bugs in the behaviour of mice or computers. As I stand at the path I have trodden with this thesis and a manuscript as my trail, I am incredibly proud of what I have achieved, and filled with gratitude for the many people who have inspired, encouraged, and supported me along the way.

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Dr. Petr Znamenskiy. Thank you for your invaluable inspiration, guidance, expertise, and support throughout this project. I deeply admire your commitment to rigour and integrity in both science and life. Your mentorship has fundamentally shaped my scientific thinking, and I will always appreciate the time and effort you have devoted to fostering my growth as a scientific researcher. I would also like to thank my thesis committee, Prof. Neil Burgess, Prof. Andreas Schaefer, and Prof. Sonja Hofer, for their valuable guidance, critical feedback and support throughout my project.

Next, I would like to extend my deepest appreciation to my lab members. Thank you Antonin, for always being a science superhero, firefighter, and pest controller (for finding and exterminating the bugs in my code). Thank you for always dedicating your time to supporting me from building electric circuits to analysing data – working with you has been a privilege. Thank you Antonio, for being a wonderful colleague, collaborator, friend, and "mouse whisperer". I sincerely appreciate your support for all the late night experiments and feeding the insatiable mice. Thank you Alex, my fellow traveller on this PhD journey, for your encouragement and friendship to make this challenging path more manageable and meaningful. Thank you Benita and Caroline, for your continuous support and for sharing the weekend mouse duties. I would like to thank all the researchers and friends in the Neurophys community at the Crick. Gaia, Mingran, Alex, Yuxin, Francesco, Irene and all my friends – you have made the past four years colourful and enjoyable with your friendship, encouragement and inspiration.

I would also like to extend my sincere gratitude to all the scientists who have supported me throughout my academic journey. I am deeply grateful to my previous supervisors at Oxford and UCL, Dr. Yanfeng Zhang, Dr. Helen Barron, Prof. Stephanie Cragg, Prof.

David Dupret and Prof. Thomas Wills, for providing invaluable guidance and laying the foundation for my academic growth. A special thank you to my TENSS family, especially Bruno, Fred, Priyanka and Florin, for the privilege of spending three wonderful summers by the lake of Transylvania, uncovering the secrets of microscopy.

I would like to thank all the researchers who have contributed greatly to advancing our understanding of the brain and artificial intelligence. I look forward to the day when we fully crack the code of the brain.

Most importantly, I would like to express my immense gratitude to my families in China and England. Thank you, mum, dad and grandpa, for your love and unwavering support of all the choices I made in my life, and for teaching me the value of perseverance, curiosity and humility. Thank you, grandpa, for always being my role model and inspiring me with the love of science and mankind. Thank you, Connor, for always being there for me during the ups and downs in my life. I am extremely lucky to go through this journey with such a kind and intelligent companion. Thank you, Mark and Delia, for making England another place I can call home. I love you all very much, and this accomplishment is as much yours as it is mine. I would also like to thank my lifelong friends, Hanzhen, Yudan, Yijun, Sunny, Amalie, and all my friends - thank you for always having my back.

Reflecting on my path so far, it has been anything but smooth or straightforward. From studying anthropology, I gradually discovered my passion for the brain, and later, artificial ones. As I take my next steps, I will challenge myself in a new path, uncovering the rationalities and irrationalities of the business world. I want to close with one of my favourite quotes from J.R.R. Tolkien: *"All that is gold does not glitter, not all those who wander are lost."*

UCL Research Paper Declaration

UCL Research Paper Declaration Form referencing the doctoral candidate's own published work(s)

Please use this form to declare if parts of your thesis are already available in another format, e.g. if data, text, or figures:

- have been uploaded to a preprint server
- are in submission to a peer-reviewed publication
- have been published in a peer-reviewed publication, e.g. journal, textbook.

This form should be completed as many times as necessary. For instance, if you have seven thesis chapters, two of which containing material that has already been published, you would complete this form twice.

1. For a research manuscript that has already been published (if not yet published, please skip to section 2)

- a) What is the title of the manuscript?
- b) Please include a link to or doi for the work
- c) Where was the work published?
- d) Who published the work? (e.g. OUP)
- e) When was the work published?
- f) List the manuscript's authors in the order they appear on the publication
- g) Was the work peer reviewed?
- h) Have you retained the copyright?
- i) Was an earlier form of the manuscript uploaded to a preprint server? (e.g. medRxiv). If 'Yes', please give a link or doi)

If 'No', please seek permission from the relevant publisher and check the box next to the below statement:

☐

I acknowledge permission of the publisher named under **1d** to include in this thesis portions of the publication named as included in **1c**.

2. For a research manuscript prepared for publication but that has not yet been published (if already published, please skip to section 3)

- a) What is the current title of the manuscript?

A depth map of visual space in the primary visual cortex

- b) Has the manuscript been uploaded to a preprint server? (e.g. medRxiv; if 'Yes', please give a link or doi)

Yes, <https://doi.org/10.1101/2024.09.27.615442>

- c) Where is the work intended to be published? (e.g. journal names)

Science

- d) List the manuscript's authors in the intended authorship order

Yiran He, Antonio Colas Nieto, Antonin Blot, Petr Znamenskiy

- e) Stage of publication (e.g. in submission)

In revision for resubmission

3. For multi-authored work, please give a statement of contribution covering all authors (if single-author, please skip to section 4)

Yiran He: developed the two-photon recording and virtual reality setup; performed the majority of the two-photon recording experiments and data analysis; wrote the manuscript in collaboration with Petr Znamenskiy.

Antonio Colas Nieto: contributed to the experiments mapping the distribution of depth representation across V1.

Antonin Blot: supervised and assisted the building of the two-photon recording and virtual reality setup; performed analysis on the detection of monitor frames and the extraction of eye tracking traces; provided guidance on data analysis.

Petr Znamenskiy: supervised the project and provided guidance on setup development, experimental design and data analysis; contributed to the analysis on three-dimensional receptive fields of V1 neurons and the distribution of depth representation across V1; wrote and revised the manuscript in collaboration with Yiran He.

4. In which chapter(s) of your thesis can this material be found?

Chapter 1, 2, 4, 5, 6, 7.

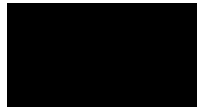
5. e-Signatures confirming that the information above is accurate (this form should be co-signed by the supervisor/ senior author unless this is not appropriate, e.g. if the paper was a single-author work)

Candidate



Date: 19/01/2025

Supervisor/ Senior Author (where appropriate)



Date 20/01/2025

Contents

Table of Acronyms	14
List of Figures	15
1 Introduction	17
1.1 Depth perception	18
1.1.1 Visual and oculomotor cues for depth perception	18
1.1.2 Behavioural studies of depth perception	26
1.1.3 Nature and nurture of depth perception	27
1.2 The mouse visual system	29
1.2.1 Visual pathways in mice	29
1.2.2 Organisation and connectivity of the mouse visual cortex	31
1.2.3 Tuning properties of neurons in the mouse visual cortex	32
1.2.4 Locomotion-related modulation of neurons in the mouse visual cortex	35
1.3 Neural mechanisms of depth perception	38
1.3.1 Neural mechanisms of depth perception from binocular disparity . .	38
1.3.2 Neural mechanisms of depth perception from motion parallax . . .	42
1.3.3 Neural mechanisms of depth perception from pictorial cues	45
1.4 Project overview	46
2 Methods	48
2.1 Authorship declaration	48
2.2 Animal welfare compliance	48
2.3 Examining the innate nature of depth perception using the visual cliff test	48
2.3.1 Visual cliff	48
2.3.2 Dark rearing	50
2.3.3 Visual cliff with optogenetic inhibition	50
2.4 Characterising the neural representation of depth in the primary visual cortex	51
2.4.1 Animals	51
2.4.2 Cranial window implantation	51
2.4.3 Virus injection	52
2.4.4 Virtual reality setup and visual stimuli	52
2.4.5 Calibration of screen positions	54

2.4.6	Eye camera calibration	55
2.4.7	Behavioral training	55
2.4.8	Two-photon calcium imaging	55
2.4.9	Widefield calcium imaging	56
2.4.10	Grating stimuli	56
2.4.11	Data analysis	57
3	Depth perception in mice does not require prior visual experience	67
3.1	Authorship declaration	67
3.2	Background	67
3.3	Mice can discriminate depth using visual cues on the customised visual cliff setup	70
3.4	Mice can discriminate depth on the visual cliff with limited pictorial cues .	71
3.5	Texture density does not bias the decision of mice on the visual cliff	73
3.6	Depth perception in mice does not depend on prior visual experience . . .	74
3.7	Superior colliculus (SC) may not be required for depth perception in mice .	77
3.8	Discussion	79
3.8.1	The nurture and nature of depth perception	79
3.8.2	Neural substrates for depth perception	80
3.8.3	Limitations	81
3.8.4	Summary	82
4	The representation of depth from motion parallax in the primary visual cortex of head-fixed mice	83
4.1	Authorship declaration	83
4.2	Background	83
4.3	Manipulating virtual depth using motion parallax in head-fixed mice . . .	86
4.4	L2/3 excitatory neurons in V1 are tuned to depth from motion parallax . .	92
4.5	Depth selectivity is invariant to stimulus size	99
4.6	Discussion	103
4.6.1	Studying the neural representation of depth from motion parallax using virtual reality	103
4.6.2	The representation of depth in V1	104
4.6.3	Summary	105
5	The emergence of depth-selective activity from the integration of loco- motion and visual motion signals	106
5.1	Authorship declaration	106
5.2	Background	106
5.3	Depth selectivity arises from conjunctive coding of running speed and optic flow speed	107

5.4	Preferred depth is determined by the ratio of preferred running speed and preferred optic flow speed	114
5.5	Closed loop coupling between locomotion and optic flow enables accurate representation of depth	119
5.6	Speed preferences of V1 neurons are consistent across conditions of passive viewing and self-generated visual flow	124
5.7	Discussion	128
5.7.1	Locomotion-related modulation supports depth representation in V1	128
5.7.2	Predictive coding	129
5.7.3	Circuit mechanisms of depth representation in V1	131
5.7.4	Closed loop coupling is required for the accurate representation of depth	132
5.7.5	Summary	133
6	A depth map of the visual space from motion parallax in the mouse primary visual cortex	134
6.1	Authorship declaration	134
6.2	Background	134
6.3	V1 neurons have three-dimensional receptive fields	135
6.4	The representation of depth across V1	137
6.5	Discussion	141
6.5.1	A depth map of the visual space in V1	141
6.5.2	The distribution of depth representation across V1	142
6.6	Summary	143
7	Discussion	144
7.1	Summary	144
7.2	Depth perception in naturalistic environments	146
7.2.1	Object motion	147
7.2.2	Free navigation	148
7.2.3	Naturalistic scenes	149
7.3	A depth map in the visual cortex	149
7.3.1	Depth representation across the visual cortex	149
7.3.2	A novel perspective of V1	150
7.3.3	A comprehensive depth map in the visual cortex	151
7.3.4	The role of V1 in depth perception from motion parallax across species	153
7.4	Functions of locomotion-related modulation in V1	153
7.4.1	Cortical state modulation and enhanced encoding of visual information	154

7.4.2	Predictive coding	155
7.5	The role of V1 depth selectivity in spatial perception	157
7.6	Circuit mechanisms of depth coding in mouse V1	160
7.6.1	Long-range inputs of locomotion signals	160
7.6.2	Local circuitry for visuomotor integration	161
7.7	Conclusion and future outlooks	163

Table of Acronyms

Notation	Description
2D	Two-dimensional.
3D	Three-dimensional.
cpd	Cycles per degree.
DMS	Dorsomedial striatum.
DSI	Direction selectivity index.
HVA	Higher visual areas.
L2/3	Layer 2/3 (of the neocortex).
L4	Layer 4 (of the neocortex).
L5	Layer 5 (of the neocortex).
L6	Layer 6 (of the neocortex).
LD	Laterodorsal nucleus in the thalamus.
LEC	Lateral entorhinal cortex).
LP	Lateral posterior nucleus in the thalamus.
IT	Inferior temporal cortex.
M2	Secondary motor cortex.
MEC	Medial entorhinal cortex.
MT	Medial temporal cortex.
POR	Postrhinal cortex.
PPC	Posterior parietal cortex.
PV	Parvalbumin.
PVP	Position-vestibular-pause.
RF	Receptive field.
RSC	Retrosplenial cortex.
SC	Superior colliculus.
SST	Somastatin.
SVM	Support vector machine.
V1	Primary visual cortex.
VIP	Vasoactive intestinal peptide.

List of Figures

3.1	Mice can discriminate depth using visual cues on the customised visual cliff setup.	71
3.2	Mice can discriminate depth on the visual cliff with limited pictorial cues. .	72
3.3	Texture density does not bias the decision of mice on the visual cliff.	73
3.4	Depth perception in mice does not depend on visual experience.	76
3.5	Superior colliculus may not be required for depth perception in mice.	78
4.1	Manipulating virtual depth using motion parallax in head-fixed mice.	87
4.2	Closed loop synchronisation and latency.	89
4.3	Screen and camera calibration using ArUco markers.	91
4.4	Quantifying depth selectivity for an example neuron.	93
4.5	Depth selectivity is widespread among V1 L2/3 excitatory neurons.	94
4.6	Depth-selective neurons measured by different calcium indicators.	96
4.7	Depth selectivity in V1 is driven by motion parallax.	98
4.8	Running, optic flow speeds and eye movements across different virtual depths.	100
4.9	Virtual depth selectivity is invariant to stimulus size.	102
5.1	Activity of depth-selective neurons as a function of running and optic flow speeds.	109
5.2	Example running speed and optic flow speed traces and and example fitted neuronal activity of the 5 models.	111
5.3	Depth selectivity arises from conjunctive coding of optic flow speed and running speed.	113
5.4	The relationship between preferred depth, preferred running speed and preferred optic flow speed.	115
5.5	Preferred depth is determined by the ratio of preferred running speed and preferred optic flow speed.	118
5.6	Open loop stimuli breaks the coupling between locomotion and optic flow feedback.	119
5.7	Conjunctive coding of locomotion and optic flow speeds does not rely on their closed loop coupling.	121
5.8	Closed loop coupling between locomotion and optic flow enables accurate representation of depth by neuronal populations in V1.	123

5.9	Spatiotemporal tuning of example V1 neurons.	126
5.10	Speed preferences mapped by three-dimensional stimuli and two-dimensional gratings.	127
5.11	Model of depth representation in the visual cortex.	130
6.1	Schematics of linear receptive field model.	136
6.2	Depth-selective neurons have three-dimensional receptive fields.	137
6.3	Depth preferences display a salt-and-pepper distribution.	138
6.4	The representation of depth across V1.	139
6.5	Relationship between depth preferences and RF locations.	140
7.1	Model of the transformation from retinocentric to egocentric representation.	158

Chapter 1

Introduction

All animals live, navigate and forage in a three-dimensional environment. Understanding the three-dimensional structure of the surrounding visual scene is vital for animals to perform daily tasks. To guide behaviour in such environment using vision, the visual system of the brain needs to infer the missing depth information from the two-dimensional projections of the environment on the retinae. The significance of depth perception is highlighted by the fact that many animals possess an innate ability to perceive depth (Walk & Gibson, 1961). Nonetheless, how the visual system parses the structure of three-dimensional scenes to estimate depth remains poorly understood. In particular, the neural mechanisms underlying an important dynamic depth cue – motion parallax – is largely unexplored (Kim et al., 2016). Therefore, this thesis aims to investigate the neural mechanisms underlying the representation of depth from motion parallax by the visual system. Specifically, I will use mice as the model organism. The mouse visual system has emerged as a powerful model to study the neural mechanisms underlying visually-guided behaviours (Niell, 2015; Niell & Scanziani, 2021). The mouse visual system has been thoroughly characterised in terms of neuronal responses to visual and non-visual stimuli (Niell & Scanziani, 2021; Niell & Stryker, 2008, 2010; Saleem et al., 2013). Studying visual system via mice allows us to harness readily available genetic toolbox and invasive recording techniques to manipulate and record activity of specific neuronal populations in this well-characterised visual system (Luo et al., 2008; Niell, 2015; Niell & Scanziani, 2021; O'Connor et al., 2009). The rapid life cycles of mice (Dean & Nachman, 2009) allows faster and larger quantity of data collection. Moreover, the pronounced resemblance in the fundamental properties and circuit motifs of the visual system between mouse and primates (Lien & Scanziani, 2018; Liu et al., 2011; Niell & Scanziani, 2021; Niell & Stryker, 2008; Priebe & Ferster, 2008; Van den Bergh et al., 2010) enables us to extrapolate our findings in mice to other higher mammals.

In this chapter, I will first provide an overview of the importance of depth perception and various visual cues that can facilitate depth perception. Then, I will outline the main components of the visual system in mice and particularly focus on the neuronal properties in the visual cortex. Lastly, I will review the existing studies on the neural mechanisms

of depth perception.

1.1 Depth perception

The perception of depth is an essential computation that animals must perform to live in this three-dimensional world. To guide behavioural tasks such as navigation, foraging, prey detection and avoiding predators using vision, animals must understand the distances of themselves to target objects and construct an internal representation of the three-dimensional structure of the environment. The ability to estimate depth has been observed across many species in the animal kingdom, ranging from insects, crustaceans, fish, amphibians, reptiles, birds to mammals (Howard, 2012c). How the perception of depth is formed has been a long-standing subject of investigation and debate in neuroscience. Helmholtz (1868) has pointed out how depth perception must be an important computation performed by the brain in one of his lectures (Bishop & Pettigrew, 1986): *"We come now to the question: how is it possible for two flat images on the retina, each from a different perspective and each representing only two dimensions, to combine so as to present a single, solid image of three dimensions? ... The combination of these sensations into a single perceptual image of the external world ... must therefore be produced, not by any anatomical mechanism of sensation, but by a mental act."*

1.1.1 Visual and oculomotor cues for depth perception

Exactly as how Helmholtz (1868) pointed out, the visual system of the brain must deduce the missing depth information based on various cues in the two-dimensional projections of the environment on the retinae. Depth perception is facilitated by various visual and oculomotor cues that utilise binocular and monocular information (Howard & Rogers, 2002). Binocular cues include two main types, binocular disparity and vergence (Howard & Rogers, 1995). Humans, primates and carnivores with front-facing eyes have extensive binocular overlap (120° in humans, Read, 2021; Spector, 1990), making binocular vision particularly important for depth perception in these species. Monocular cues include accommodation, pictorial cues originated from static images, and motion parallax and other motion-based cues created from relative motion between the observer and the environment (Howard, 2012a, 2012b). Most animals do not possess frontal eyes or a binocular overlap as wide as that of primates (Read, 2021). For example, mice only possess 40° of binocular vision with an extensive periphery vision (Heesy, 2004). Therefore, monocular cues play an important role in the depth perception of these species.

Binocular disparity and vergence

Binocular vision requires integration of images from both eyes. The horizontal separation of the two eyes results in slight offsets between two monocular images, and two

monocular views are integrated through binocular fusion, producing a single, coherent visual percept of the environment (Ambrad Giovannetti & Rancz, 2024; Ding & Levi, 2021). In primates and humans, vergence eye movements (the coordinated convergence or divergence of eyes) align the foveae to fixate eyes onto a specific target, contributing to the motor aspect of binocular fusion (Harwerth et al., 1996; Miles, 1997; Schor et al., 2002). The fixation point, along with all spatial locations that project to corresponding retinal locations sharing single subjective locations in the visual field, define the horopter; objects on the horopter produce zero disparity between two monocular images (Turski, 2023). Objects outside the horopter project to non-corresponding retinal points and hence generate non-zero binocular disparity (Howard & Rogers, 1995; Howard & Rogers, 2002; Qian, 1997). However, if objects are situated within a certain range to the horopter (Panum’s fusional area), the visual system can still fuse the slight offset of monocular objects into a single percept of the object (Poggio & Poggio, 1984; Turski, 2023). If disparities exceed this range, sensory fusion fails and diplopia (double vision) occurs. Further vergence movements are required to adjust the alignment between two eyes and bring the object into Panum’s area to allow binocular fusion (Poggio & Poggio, 1984).

Both vergence and binocular disparity can serve as binocular cues for depth perception. When eyes fixate on an object at the middle plane between two eyes, the absolute distance of the object to the observer can be calculated from the vergence angle θ and the interocular distance I by $\frac{I}{2 \tan \theta/2}$ (Howard, 2012a). Convergent movements of two eyes to fixate on a different object indicates a nearer depth, whereas divergence signifies a farther depth. Vergence and binocular disparity can be combined to determine the absolute distance of an object to an observer. If the observer focuses their eyes at a certain point with a vergence angle of θ , another objects outside the horopter generate a binocular disparity of σ with respect to the fixation point, depth can be calculated by $\frac{I}{2 \tan \frac{\theta - \sigma}{2}}$ (Howard, 2012a). However, there is a limited range over which binocular cues effectively supports depth perception. Various studies suggest that for humans, binocular disparity is most effective for depth judgements up to several meters as disparity decreases with distance (around 6 to 10 m, although some suggest that binocular cues can significantly improve depth perception when monocular cues are weak up to 30 m, Blakemore, 1970; Crannell and Peters, 1970; Hirsch and Weymouth, 1948; Howard, 1919; Sweet and Kaiser, 2011). Similarly, as vergence angles barely change beyond 2 m, it is unlikely that vergence provides a primary cue for depth beyond that distance (Howard, 2012a).

Binocular depth perception does not necessarily require vergence or even binocular fusion. For humans, Dove (1841) showed that stereoscopic vision still exists with a brief flash of light, too brief for vergence to occur. Fine disparity arises within Panum’s fusional area, enabling more precise depth estimation in the central visual field (Turski, 2023). Outside fusible range and even with double vision, coarse disparity can still provide a sense of stereoptic depth and is especially important for creating sense of immersion in the ambient environment (Ogle, 1953; Turski, 2023). For other animals such as rodents,

they do not fixate on specific objects in the same way as primates (Stahl, 2004). Mice, for instance, have laterally oriented eyes that diverge at $\sim 103^\circ$ with only $\sim 40^\circ$ of binocular overlap (Heesy, 2004). Their eye movements primarily serve as compensation to head movements, such as tilt and rotation (Holmgren et al., 2021; Meyer et al., 2020; Michaeli et al., 2020). Mice can discriminate depth based on binocular disparity cues even if their vergence eye movements are not correlated with disparity (Choi & Priebe, 2020; Samonds et al., 2019).

Binocular vision facilitates depth judgements in high precision and accuracy (Kim et al., 2016; McKee & Taylor, 2010). Humans can make much more precise judgements about the relative distance between two objects using binocular vision than using monocular vision in both simple and enriched settings (McKee & Taylor, 2010). However, binocular vision is not necessarily required for depth perception. Most animals do not possess frontal eyes or a binocular overlap as wide as that of primates (Read, 2021). Human infants and rodents with only one eye displayed unimpaired ability to perform depth perception tasks (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022; Walk & Dodge, 1962; Walk & Gibson, 1961), underscoring the crucial role of monocular visual information for animals' depth perception.

Accommodation

There are rich monocular oculomotor and visual cues in the environment for depth perception. Accommodation is an monocular oculomotor depth cue that signals depth by changing the shape of the lens of the eyes or order to bring the target object in focus (Berkeley, 1709; Descartes, 1664; Howard, 2012b). Nearby objects need higher convexity of the lens while farther objects requires lower convexity. Humans can use accommodation alone as a cue for perceiving depth. When presenting subjects with a black silk thread or the edge of a black card, subjects were able to detect the changes in depth of several centimeters based on changes in accommodation (Hillebrand, 1894; Wundt, 1862). Regarding estimating absolute depth from accommodation, (Fisher & Ciuffreda, 1988) found that although most people could give a decent estimation of depth when viewing high-contrast stimuli at different distances, only a quarter of the participants show strong correlation between the perceived depth and the accommodative response of their eyes. This indicates that at least for some people, accommodation alone can be a useful cue for depth. Blur of objects that are out of focus can also provide depth information. When blur was isolated through a controlled optical setup as the only cue for depth, subjects were able to report the depth a few centimetres from the actual depth (Grant, 1942). Humans can also distinguish the image closer than the focal plane and that beyond the focal plane (Wilson et al., 2002). Accommodation is most effectively at distances from several centimetres up to around 2 m, due to optical power limits at closer range and minimal accommodative changes at farther distances (Chen et al., 2000; Howard, 2012a).

Pictorial cues

Monocular depth cues include pictorial cues that indicate depth from single static images, and motion-based cues that provides depth information based on the relative movement between the observer and objects in the visual scene (Howard & Rogers, 1995; Howard & Rogers, 2002; Kim et al., 2016; Reichelt et al., 2010; Tovée, 2008). There are several types of pictorial cues for depth estimation (reviewed in Daw, 2012). (1) Linear perspective: the famous artists, Leon Battista Alberti and Leonardo da Vinci, described linear perspective as the phenomenon where parallel lines appear to meet at a single vanishing point in the distance, creating the illusion of depth on a flat surface. (2) Relative size: objects will appear smaller at a further distance, and the relative sizes of objects can be used to estimate relative depth as well as absolute depth for objects with known sizes. (3) Texture gradients: objects at various distances form texture gradients, where the density and clarity of textures change with distance. (4) Aerial perspective: the atmosphere in front of objects and the scattering of light makes objects further appear hazy and bluer, creating aerial perspective. (5) Occlusion: occlusion and superposition provide information about an object in front of another object. (6) Shading: the shading of an object will change depending on its relative distance to the light source and its shape.

Motion parallax

The sensation of depth can be also generated from the relative motion between the observer and objects in the environment. One important and powerful motion-based cue is motion parallax. First described by Helmholtz (1925) as a potential source of three-dimensional information, motion parallax usually refers to the pattern of objects moving across the retina due to translational movement of the observer's viewpoint (observer-produced parallax) or the objects (object-produced parallax, Gibson et al., 1959; Nawrot and Joyce, 2006; Rogers and Graham, 1979. When an observer translates in the environment relative to the objects, objects at different distances appear to move at different velocities on the retina Gibson et al., 1959; Helmholtz, 1925; Rogers & Graham, 1979. The translation of the observer in the environment creates optic flow emerging from a focus of expansion and disappearing into a focus of contraction Miles, 1998. If the observer fixates at optical infinity or do not fixate like rodents, all parts in the retinal image will move in the direction opposite to the self-motion direction of the observer, and the speed of visual motion on the retina is inversely proportional to the distance Miles, 1998; Yoonessi & Baker, 2011. If the observer fixates at an intermediate distance, objects closer than the fixation point will move in the direction opposite to the observer's movement, whereas objects farther than the fixation point will move in the same direction as the observer's movement Miles, 1998; Yoonessi & Baker, 2011. The visual motion speed is inversely proportion to the distance between the object and the plane of fixation Miles, 1998. Pure eye rotation or pure head and body rotation along the vertical axis of the

body (if ignoring the translation caused by the offset between eye positions and the axis of body rotation) creates optic flow patterns resembling the latitude of a sphere, yet the changes in visual angle do not differ as a function of distance Miles, 1998.

Based on the rules of motion parallax, the absolute depth of a stationary object in an environment can be computed if we know the visual motion speed of the image on the retina and the speed of changes in gaze direction as a result of translation or rotation of eye, head and body (Howard, 2012b). If an object is located at an angle α from the visual axis of the observer, and the observer is translating at a speed of dr leading to the retinal image of the objects moving at a speed of $d\theta$, the distance of the object from the observer is proportional to the ratio of running speed and optic flow speed ($\sin \alpha \frac{dr}{d\theta}$). However in practice, ambiguity in the perception of absolute depth could arise from incorrect perception of the motion of the objects (whether the object is stationary or moving at a certain speed and whether the motion is purely translational) or the inaccurate perception of the motion of the observer's eyes, head and body. This may cause visual illusions such as stationary objects such as a printed face or a stereogram appear to move when we observe them with lateral head movements (Howard, 2012b; Shimono et al., 2002). When we incorrectly estimate the distance of a stationary objects, the actual visual motion is different from our expectation, so we attribute the difference to object motion and create an illusion of object moving.

Regarding the perception of relative depth from motion parallax, early work have showed that humans can perceive slanting of a moving sloping surface from the velocity gradients produced at different points of the surface, although the report of the degree of relative depth or slant was inconsistent across subjects and the texture gradient of the surface was not controlled (Braunstein, 1968; Gibson et al., 1959; Gibson & Carel, 1952). Rogers and Graham (1979) later established motion parallax as an independent cue for reliable unambiguous perception of relative depth. Using random-dot stereograms, he simulated three-dimensional surfaces of different shapes and degrees of slanting, and asked subjects to report the shape and depth of the surfaces either from active translation of the observer or from the translation of the visual stimuli. Most subjects were able to report the shape and corrugation of the surfaces both under active translation or stimuli translation, and were able to match the perceived depth from motion parallax to depth from binocular disparity. The random-dot stereogram experiments eliminated all depth cues other than motion parallax, and clearly demonstrated that motion parallax alone, whether generated from self-motion or external motion, is sufficient for the perception of relative depth. Notably, subjects did report a more pronounced perception of depth in the observer-produced parallax condition, signifying the importance of self-motion signals in depth perception. Similar to humans, non-human primates such as rhesus monkeys can discriminate relative depth in random-dot stereograms from motion parallax cues created by movement of dots or passive translation of the monkeys (Cao & Schiller, 2002; Kim et al., 2015a).

Most animals do not have a binocular overlap as extensive as primates (Read, 2021), and animals such as rodents cannot fixate their eyes on objects in the same way as primates (Stahl, 2004). Therefore, monocular depth cues can provide a major source of depth information for these animals. For examples, Mongolian gerbils and mice were found to display frequent head bobs (vertical head movements) before jumping across a gap, which were proposed to create motion parallax cues to assist depth estimation (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022). In gerbils, the number and frequency of head bobbing were positively correlated with the gap distance the animal was required to jump, and the accuracy of jumping measured by the variance of landing positions were enhanced with more frequent large head bobs. Under the condition where looming cues were limited using smaller platforms, head bobs were still correlated with the distance and accuracy of jumping. Parker, Abe, Beatie, et al. (2022) further showed that monocular mice were able to perform the jumping task as well as binocular mice, and they performed more frequent head movements when binocular cues were unavailable. The variance of landing locations was similar between binocular and monocular conditions. These results suggest that mice can accurately estimate depth in absence of binocular cues, and motion parallax is an important monocular cue for depth judgments.

In addition to terrestrial animals, there is some evidence that avian and aquatic animals may also be able to perceive depth from motion parallax cues. Insects such as bees and flies cannot rely on binocular disparity or lens accommodation as primates due to their small interocular distance, restricted binocular overlap, and the fixed focal length of lenses in their compound eyes (Evangelista et al., 2010; Horridge, 1986; Kapustjansky et al., 2010; Lehrer et al., 1988). Instead, they heavily rely on motion-related cues. Similar to head bobs in rodents, translational head movements were found in locusts when jumping towards a target (Sobel, 1990a). When presenting bees with artificial flowers at randomised size and at various heights, bees could identify the flower at the height of reward, indicating their ability of discriminating depth regardless of size (Lehrer et al., 1988). During the depth discrimination tasks, bees primarily flew in straight lines, creating translational motion parallax (Lehrer et al., 1988). When bees were trained to fly through a gap between two grating stimuli, they would fly through the centre of two stationary stimuli (Srinivasan, 1992). Yet, a moving stimulus on one side would bias the bees to fly towards the moving side, as the moving stimuli creates an illusion of nearer depth based on motion parallax, suggesting that motion parallax signals can bias the depth judgements of bees during flight (Srinivasan, 1992). A similar phenomenon has been found in flying birds (Bhagavatula et al., 2011). During landing, bees maintain constant visual motion speed of the approaching surface to ensure deceleration of their flying speed (Srinivasan et al., 2000). Flies use the same strategy to control their flying speed (David, 1982).

For aquatic animals, water flow creates global motion and presents extra challenge to estimating optic flow speed caused by self-motion. When training zebrafish to swim across

a tunnel with two grating stimuli, unlike insects, zebrafish do not maintain a constant visual flow speed by changing their swimming speed (Scholtyssek et al., 2014). Instead, they may measure optic flow speed momentarily to constantly monitor their own speed. Nevertheless, optic flow did bias the distance of zebrafish to the tunnel wall - stronger optic flow attracted the zebrafish to swim closer to that side, which was opposite to the direction of bias for bees (Scholtyssek et al., 2014). The bias in swimming towards objects with stronger optic flow might be a strategy of staying closed to structures that provide visual feedback about self-motion in low-lighting aquatic environment.

Other motion-based cues

Apart from motion parallax, the relative movements between the observer and the target object can generate other visual patterns, providing additional information for depth. For example, the image of a target object can be distorted from the perspective of an observer in relative motion with the object (dynamic perspective cues, Buckthorpe et al., 2017; Kim et al., 2015b; Rogers and Rogers, 1992). If the observer is translating their head while making compensatory eye movements to fixate on a target object, a pure eye translation creates perspective distortions in the retinal coordinates. On the contrary, a pure eye rotation results in perspective distortions on a planar projection but not in the retinal coordinates (Kim et al., 2015b). Therefore, the changes in perspective during locomotion can provide information about the observer's self-motion (translation and rotation). Without any pictorial cues or extra-retinal locomotion signals (e.g. smooth eye movement command signals), the perceived depth from vision motion alone can be ambiguous (Hayashibe, 1991; Nawrot, 2003). However, dynamic perspective cues can be used to disambiguate depth perception by the inference of the change of eye orientation relative to the environment (Kim et al., 2015b; Rogers & Rogers, 1992). When showing observers a three-dimensional corrugated surface constructed by random dots, changes in perspective of the surface alone by rotating the display around its vertical axis (which mimicked the dynamic perspective cues generated from eye rotation) can disambiguate the depth (Rogers & Rogers, 1992). However, eye movements were not recorded in these experiments, which might be an extra source of depth information. Kim et al., 2015b further provided a clearer demonstration that dynamic perspective can disambiguate depth for MT neurons in macaques. The stimuli consisted of a patch of random dots in the receptive field of a neuron and at a specific distance to the animal. In the motion parallax condition, the animal was translated along a frontoparallel plane while making compensatory eye movements to maintain fixation. In the retinal motion condition, the animal remains stationary while the stimuli moved across to mimic the retinal motion of the motion parallax condition. In the dynamic perspective condition, the small stimulus patch was surrounded by a larger cloud of random triangular dots outside the receptive field of the neuron and at various distances to the observer. The animal remained stationary while both the small stimulus patch and the larger cloud simulated the translation and

rotation of eyes in the motion parallax condition. Pure retinal motion elicited symmetrical tuning to near and far depths for MT neurons, whereas many neurons showed preferential firing to either near or far depths in the dynamic perspective condition (65%) and motion parallax condition (89%). This shows that MT neurons in macaques can disambiguate depth from retinal motion with either eye movement signal or dynamic perspective cues.

When an observer moves laterally relative to a couple of stationary objects at different distances, or when the observer remains stationary while objects at different distances are moving relative to each other, the accretion/deletion of the farther objects (dynamic occlusion) offers an unambiguous cue for relative depth (Andersen & Braunstein, 1983; Ono et al., 1988; Royden et al., 1988). Absolute depth can only be inferred when the observer is moving relative to stationary objects and the absolute distance between the observer and one of the objects is available, and then depth can be estimated based on the speed of the observer's locomotion and the rate of accretion/deletion (Ono et al., 1988). When conflicting dynamic occlusion and motion parallax cues were presented using two surfaces constructed by random dots, dynamic occlusion cues dominated the perception of depth signs in human subjects (Ono et al., 1988).

Another type of optic flow pattern is looming, referring to the increase or decrease of size when the observer moves along the visual axis towards or away from the objects (Howard, 2012b; Yoonessi & Baker, 2013). The expansion or compression of retinal images of objects as well as the looming at distinct rates for objects at different distances provide information of relative depth and motion in depth.

Cues for depth perception in mice

Mice can use a variety of binocular and monocular cues for depth estimation. Despite the relatively small binocular overlap of mice (Heesy, 2004), mice are able to use binocular disparity for depth judgements (Samonds et al., 2019). However, mice cannot fixate like primates (Stahl, 2004) and their eye movements are mainly to compensate the head movements (Holmgren et al., 2021; Meyer et al., 2020; Michaiel et al., 2020). Mice cannot change the refractive power of their eyes either (Chalupa & Williams, 2008). Therefore, vergence and accommodation are unlikely to provide main source of depth information for mice. It was indeed shown in experiments that mice can discriminate depth with binocular disparity cues whilst vergence plays little role. Samonds et al. (2019) trained mice to discriminate pairs of near and far surfaces constructed by random dots with the same magnitude of binocular disparity with respect to a monitor screen 22 cm away, and mice could indicate their judgements by running or stopping. In theory, the preferred viewing distance of mice is around 10 cm based on the optics of their eyes (Scholl et al., 2013). Based on their interocular distance of around 1 cm and their preferred viewing distance, the possible range of disparity that mice can experience is around $\pm 5.7^\circ$ (from 5 cm to infinity), which is much larger than cats ($\pm 2.5^\circ$, Packwood and Gordon, 1975) or primates ($\pm 1.5^\circ$, Blakemore, 1970). In Samonds et al. (2019)'s experiments, mice could

clearly differentiate surfaces within a range of $\pm 0.83^\circ$ to $\pm 0.5^\circ$ disparity with respect to the screen. It is worth noting the difficulty of estimating the actual disparity that mice experience, as their large depth of field complicates the estimation of their optimal viewing distance (de la Cera et al., 2006; Geng et al., 2011). However, there were no correlation between the changes in vergence angle and depth, suggesting that mice can discriminate depth based on binocular disparity without relying on vergence (Samonds et al., 2019). Boone et al. (2021) has also shown that mice were able to discriminate depth on a visual cliff test (described in more detail in the section below) by avoiding the side with a apparent visual drop, when predominantly using their upper visual field with considerable binocular overlap.

As mice enjoy a large peripheral visual field compared to primates (Heesy, 2004), monocular cues can provide a major source of depth information. As summarised in Section 1.1.1, monocular mice were able to jump across gaps of variable distances as accurately and as precisely as binocular mice, displaying more frequent head bobs to generate motion parallax cues (Parker, Abe, Beatie, et al., 2022). This suggests that motion-based monocular cues can be an important source of depth information for mice in absence of binocular cues. Pictorial cues and other motion-based cues are less studied in isolation for mice. However, many behavioural paradigms assessing depth perception such as the visual cliff test (Boone et al., 2021; Fox, 1965) and the jumping task (Parker, Abe, Beatie, et al., 2022) contain rich pictorial cues and other motion-based cues such as relative size, perspective, shading, looming, dynamic perspective, etc. For example, Ellard et al. (1984) showed that gerbils could use looming cues together with motion parallax cues to ensure accurate jumping, and mice can clearly detect looming cues with behavioural responses (Koehler et al., 2019; Yilmaz & Meister, 2013). Neurons in the mouse visual cortex are also tuned to stimulus orientation, spatio-temporal frequencies, changes in brightness (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008) and size (Adesnik et al., 2012; Dipoppa et al., 2018), which can serve as neural substrate for detecting relative size, perspective and shading. Therefore in theory, mice should be able to use pictorial cues, motion parallax and other motion-based cues to estimate depth, although further behavioural experiments will be needed to confirm this statement.

1.1.2 Behavioural studies of depth perception

As animals cannot report their perception verbally, behavioural assays are required to assess and quantify their ability to estimate depth. Earlier studies usually employed a simple approach - by putting animals at various heights and observe their reactions due to fear of heights. For example. baby pigs (Spalding, 1875) and newly hatched chicks (Thorndike, 1899) were placed at platforms of various heights and their willingness to jump down was observed. As heights increased, they displayed increased hesitation or refusal to jump down. This assay shows that for pigs and chicken, the fear of heights and

the ability to perceive a visual drop is present at the very day of birth.

Another type of behavioural paradigm stemmed from the observations of animals making horizontal or vertical head movements before jumping towards a target. Locusts make lateral head movements to perform visual scanning ("peering") before jumping towards a target, which has been proposed as a way of generating motion parallax cues from self-motion to assist more accurate depth perception (Sobel, 1990b; Wallace, 1959). When training animals to jump across a distance between two raised platforms, rats (Lashley & Russell, 1934; Russell, 1932), mice (Parker, Abe, Beatie, et al., 2022), cats, kittens (Olson, 1980), and Mongolian gerbils (Ellard et al., 1984) make vertical head movements (head bobs) prior to the jump. Their accuracy and precision of landing on the target platform as well as the frequency of head bobs were analysed as an indication of their depth perception ability (see Section 1.1.1).

In 1960s, an extensive amount of studies started to employ a behavioural assay called the visual cliff (Fox, 1965; Gibson & Walk, 1960; Walk & Gibson, 1961). This paradigm examines the ability of depth discrimination using animals' avoidance of a visual drop. The apparatus usually consists of a large transparent panel bisected by a central platform, with one side (the "shallow side") featuring a texture directly beneath the panel, whilst the other side (the "deep side") displaying another texture several feet below, creating a sensation of vertical depth. It was hypothesized that animals with a normal ability of depth discrimination would prefer descending to the shallow side due to fear of height or other possible reasons. Walk and Gibson (1961) comprehensively examined a wide range of adult animals including mammals such as rats, mice, goats, pigs, dogs and monkeys as well as birds (adult chickens). They consistently preferred to choose the shallow side or avoided the deep side, indicating a universal ability of discriminating depth across many animal species.

1.1.3 Nature and nurture of depth perception

The observations of the presence of depth perception in baby pigs and chicks at the day of birth (Spalding, 1875; Thorndike, 1899) presented an interesting question regarding the developmental origin of depth perception – is depth perception an innate ability for all animals, or does it require any prior visual experience? The visual cliff became an ideal method for assessing the innateness of depth perception due to several advantages. First, the preference to the shallow side is robust across species (Walk & Gibson, 1961). Second, unlike the jumping task (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022), it does not require prior training and high-level motor coordination for animals to perform the task. Third, depth perception can be easily quantified by measuring animals' preferences between two simple options.

To explore whether depth is innate for humans, researchers tested human infants aged 4 - 14 months on the visual cliff while the infant's mother calling them from either the deep side or the shallow side (Birch et al., 1982; Fox et al., 1980; Gibson & Walk, 1960;

Tsuruhara et al., 2014). Strikingly, most infants refused to cross the deep side even when they were encouraged by their mothers, indicating that the ability to perceive depth is present for humans even from a very young age. However, since human infants usually take several months to start crawling, it is impossible to test human infants on the visual cliff as soon as they were born. Consequently, the possibility that visual experience within the first few months helped the human infants to develop depth perception cannot be ruled out.

Precocial animals that are mobile at birth thus serve as better proxies to study whether depth perception requires prior visual experience. Chicks, lambs and kids were tested on the visual cliff and all demonstrated innate avoidance to the deep side (Gibson & Walk, 1960). Another way of abolishing prior visual experience is to raise animals completely in a dark environment. Rats born and reared in the dark for up to 4 months and then immediately tested on the visual cliff displayed a strong preference to the shallow side (Cheney & Crow, 1969; Walk et al., 1957; Walk et al., 1965). These results established depth perception as an innate ability that does not require prior visual experience for many birds and mammals.

On the other hand, for some mammals, visual experience seems important in developing or refining depth perception. For kittens, self-initiated movements accompanied with visual flow feedback is required for depth perception (Hein et al., 1970). Compared to kittens that could actively explore in an environment, kittens that could not control their own movement and were only passively exposed to visual inputs did not develop paw placing reflexes or avoid the deep side on the visual cliff (Hein et al., 1970). Infant monkeys reared in the dark for more than a month did not discriminate between the shallow and deep sides on the visual cliff (Fantz, 1965). Even rats reared in the dark for more than 140 days did not immediately show depth discrimination on the visual cliff, but required 2-4 weeks of visual experience for them to restore depth perception (Walk et al., 1965). It is possible that at least for some animals, the maturation of depth perception requires certain degrees of exposure to visual experience during a critical period of development.

Overall, depth perception is proven innate for many animals including birds (chicken) and mammals (sheep, goats and rats), whereas visual experience might be needed for some species (especially higher mammals) to develop or refine their depth perception. The innate nature of depth perception underscores its importance as a core visual function for animals to perform daily tasks. Moreover, depth perception is likely supported by evolutionarily conserved neural circuits that are largely shaped by gene expression patterns.

1.2 The mouse visual system

1.2.1 Visual pathways in mice

In mouse visual system, visual stimuli are first processed by photoreceptors including rods and cones, to convert light into electrical signals (Carter-Dawson & Lavail, 1979; Fu & Yau, 2007). Rods which are specialised in vision under low-light conditions constitute the majority of photoreceptors in the mouse retinae, and cones are specialised for vision with colour and fine details (Carter-Dawson & Lavail, 1979). Through multiple retinal interneurons including horizontal, bipolar and amacrine cells, the electrical signals were then are transmitted to the retinal ganglion cells (RGCs) as an output to the rest of the brain (reviewed in Masland, 2012; Seabrook et al., 2017). There are more than 32 types of RGCs being identified with diverse functional specialisation, each extracting specific features from the visual scene (Baden et al., 2016; Goetz et al., 2022). Different types of RGCs selectively respond to either sustained or transient light intensity increases (ON cells) or decreases (OFF cells), and specific directions, orientations or contrasts of the stimuli (Baden et al., 2016).

In the brain, RGC subtypes project to over 40 subcortical targets in the basal fore-brain, hypothalamus, thalamus, superior colliculus, pretectum, and other midbrain regions (Morin & Studholme, 2014). These projections provide visual information for either image-forming circuits directly leading to visual perception or non-image-forming circuits for other types of perception or eye-related reflexes to support vision (Morin & Studholme, 2014). Areas in the image-forming visual pathway generally retain a spatially organised representation of the retinal image, known as retinotopic maps, derived from RGC inputs. This means that individual neurons in these regions respond to specific locations in the visual field and neighbouring neurons respond to adjacent areas in the visual field (Seabrook et al., 2017). There are also RGCs, such as melanopsin-containing intrinsically photosensitive RGCs (ipRGCs), that provide inputs to the non-image-forming pathway (Beier et al., 2020; Ecker et al., 2010; Güler et al., 2008; Panda et al., 2002; Schmidt & Kofuji, 2008). Projecting to suprachiasmatic nucleus (SCN) in the hypothalamus and olivary pretectal nucleus (OPN) in the midbrain (Hattar et al., 2002), ipRGCs were found to drive circadian clock resetting and pupillary light reflex through their responses to light (Güler et al., 2008; Hatori et al., 2008; Lucas et al., 2003; Panda et al., 2002). A recent study has also found that the projections from GABAergic ipRGCs to the ventral lateral geniculate nucleus (vLGN) in the thalamus could also be involved in innate fear responses to heights (Shang et al., 2024).

There are two main pathways that transfer retinal information to the visual cortex – the primary site for conscious visual perception: first, the direct relay of visual information from RGCs via the dorsal lateral geniculate nucleus (dLGN), or second, the indirect route from RGCs via the superior colliculus (SC) and other intermediate stations including dLGN, the lateral posterior nucleus (LP) and the lateral dorsal nucleus (LD) (Ellis et

al., 2016; Liu et al., 2022; Martin, 1986; Román Rosón et al., 2019; Seabrook et al., 2017). dLGN is a key relay station of retina information in the image-forming pathway, with 30 - 40% of RGCs projecting to dLGN (Martin, 1986). The majority of RGC axons cross at the optic chiasm to the contralateral side (Jaubert-Miazza et al., 2005). The small amount of ipsilateral RGC axons are localised to the medial dLGN, while the contralateral axons occupy a wider non-overlapping area of the dLGN (Reese, 1988; Reese & Jeffery, 1983). Axons from the nasal hemiretina cross to the contralateral side while axons from the temporal hemiretina remain on the ipsilateral side, leading to the left dLGN receive information from the right visual field and the right dLGN from the left visual field from both eyes (Wurtz & Kandel, 2000). Axons from neighbouring RGCs project to adjacent regions in the dLGN, creating retinotopic maps (McLaughlin & O’Leary, 2005; Reese, 1988; Reese & Jeffery, 1983; Sitko et al., 2018). dLGN then in turn projects to layer 4 (L4) of the primary visual cortex, providing the major source of bottom-up visual input (Gilbert & Wiesel, 1983).

SC is another important midbrain structure that receives the majority of RGC projections and provides visual information for the non-image-forming pathway (Ellis et al., 2016). SC is an evolutionarily-conserved, integrative sensorimotor structure receiving inputs from multiple sensory modalities, and is involved in controlling innate behaviours such as eye movements, head movements and escaping from looming objects (Ito & Feldheim, 2018; Shang et al., 2015; Sparks et al., 1990; Wei et al., 2015). Organised in layers with distinct innervation, its superficial layers receive input from over 85% of retinal ganglion cells (Ellis et al., 2016), and its deeper layers receive projections from the superficial layers as well as the primary motor, somatosensory and auditory cortices (Dräger & Hubel, 1976; Ito & Feldheim, 2018; Triplett et al., 2012; Zingg et al., 2017). It indirectly conveys information to layer 1 (L1) of the visual cortex via its projections to thalamic nuclei including dLGN and LP (Bickford et al., 2015; Gale & Murphy, 2014; Roth et al., 2016; Seabrook et al., 2017; Tohmi et al., 2014). It has been well-established that the deep layers of SC is involved in controlling saccadic eye movements for stabilisation of retinal images (Straschill & Hoffman, 1969; Wang et al., 2015; Wurtz & Goldberg, 1972). The superficial SC layers contain a map of the contralateral retina, with the dorsal-ventral and temporal-nasal axes of retina mapping onto the lateralmedial and anteriorposterior (AP) axes of the SC respectively (Cang & Feldheim, 2013; Dräger & Hubel, 1976; Triplett, 2014). The small amount of ipsilateral RGC axons originating from the ventrotemporal retina project to deeper SC layers (Dräger, 1975), and the deeper SC layers contain auditory, somatosensory, and eye movement maps that are aligned with the retinotopy of the superficial SC (Cang & Feldheim, 2013; Wang et al., 2015). The projections from SC to LP and the outer shell of dLGN also maintain the retinotopic structure (Bickford et al., 2015; Roth et al., 2016).

1.2.2 Organisation and connectivity of the mouse visual cortex

The mouse visual cortex consists of V1 and several higher visual areas (HVAs), including the lateromedial (LM), the anterolateral (AL), the rostrolateral (RL), the anterior (A), the anteromedial (AM), the posteromedial (PM), the laterointermediate (LI), the posterior (P) and the postrhinal (POR) (Garrett et al., 2014; Glickfeld & Olsen, 2017; Wang & Burkhalter, 2007; Wang, Ding, et al., 2020). V1 and HVAs send strong feedforward and feedback projections to each other (DSouza et al., 2022; Glickfeld & Olsen, 2017; Morimoto et al., 2021; Nurminen et al., 2018; Olavarria & Montero, 1989; Vangeneugden et al., 2019; Wang & Burkhalter, 2007). HVAs are also interconnected, with each HVA directly connecting to all other HVAs with diverse connection densities (Wang et al., 2011, 2012). It is believed that while V1 is responsible for low-level processing and extraction of visual features, HVAs are capable of higher-order processing to facilitate complex visually-guided tasks (Glickfeld & Olsen, 2017). Different HVAs are thought to specialise in distinct functions. It has been proposed that HVAs in mouse can be divided into the ventral and dorsal streams analogous to primate visual cortex (Kravitz et al., 2011), with the medial/anterior areas (AL, RL, A, AM, PM) preferentially connected to parietal, motor and limbic cortices for spatial perception (the "where" pathway), and lateral areas (LM, P, LI, POR) more strongly linked to temporal and parahippocampal cortices for object perception (the "what" pathway Wang et al., 2012). V1 inputs to HVAs show similar trend of specialisation – distinct V1 populations were found to project to AL and PM (Han et al., 2018; Kim et al., 2018).

Mouse V1 is comprised a monocular zone and a small binocular zone (Dräger, 1975; Kalatsky & Stryker, 2003; Schuett et al., 2002; Wagor et al., 1980). The monocular zone maps to the contralateral visual field, and the binocular zone receives projections from both the contralateral and ipsilateral eye (Dräger and Olsen, 1980; Herrera et al., 2003; Petros et al., 2008), mapping to the binocular overlap between two eyes. Mouse V1 and HVAs inherit the retinotopic organisation of the subcortical pathway (Garrett et al., 2014). While V1 representation covers the whole visual field, HVA representations are biased towards particular parts of the visual field, indicating a more specialised processing of certain visual information (Garrett et al., 2014). For example, RL is biased towards the lower nasal visual field, PM predominantly covers the temporal visual field and P prefers the upper visual field.

V1 receives projections from various cortical and subcortical areas which convey rich information beyond pure visual signals. These projections include the secondary motor cortex (M2), the hippocampal complex, the retrosplenial cortex (RSC), and multiple thalamic nuclei such as LGN, LP and LD (Froudarakis et al., 2019; Juavinett et al., 2020). M2, RSC and the hippocampal complex convey rich signals encoding locomotion, head direction, eye movement and spatial location (Leinweber et al., 2017; Saleem et al., 2018; Vélez-Fort et al., 2018). Apart from relaying retinal visual information, dLGN and LP both convey locomotion signals to V1, with dLGN signals positive combination of

running speed and visual motion speed and LP responding to the discrepancies between self-motion and visual motion (Roth et al., 2016).

Regarding laminar circuitry, the main feedforward circuit of visual information in the visual cortex generally follows the canonical cortical microcircuits as other sensory cortices (Adesnik & Naka, 2018; Gilbert & Wiesel, 1983; Harris & Mrsic-Flogel, 2013; Petersen & Crochet, 2013). Although all layers of the visual cortex receive thalamic projections, layer 4 (L4) receives the strongest innervation (Ji et al., 2016). Next, thalamic inputs are relayed to layer 2/3 (L2/3) and then convey the signals to deeper layers layer 5 (L5) and 6 (L6) as the main cortical output to subcortical structures. As the major input layer, L4 mainly integrates signals from multiple dLGN neurons, and is proposed to conduct initial processing of visual information (reviewed in Niell and Scanziani, 2021). L2/3 excitatory neurons link the input and output layers of the visual cortex and also send projections throughout the neocortex, and thus are considered to play a critical role in integration of cortical information (Harris & Shepherd, 2015).

1.2.3 Tuning properties of neurons in the mouse visual cortex

Following the introduction of "receptive fields" (RFs) by Sherrington (1906), visual neuroscientists use RFs to describe a specific region of the visual field where the presence of visual stimuli can trigger a response of the neurons (Hartline, 1938; Spillmann, 2014; Wandell, 1995; Wandell & Winawer, 2015). Early studies found that RFs of RGCs follow a circular center-surround pattern – the ON-centre cells which are excited by light in the centre and inhibited by light in the surround, and the OFF-centre cells which are excited by the opposite light pattern (Barlow, 1953; Hartline, 1938, 1940; Kuffler, 1953). RFs of dLGN relay cells measured in cats are similar to the circular and antagonistic center-surround organisation of their retinal inputs (Hubel & Wiesel, 1961). However, RFs of neurons in the visual cortex are drastically different from the circular and concentric RFs of dLGN neurons and RGCs. RFs of cells in the cat visual cortex constitute of 2 categories – simple cells and complex cells (Hubel & Wiesel, 1962). Simple cell RFs consist of elongated ON and OFF subregions which can be modelled by Gabor functions, and their responses can be predicted by a linear summation of responses to the input image (Bonin et al., 2011; Hubel & Wiesel, 1962; Niell & Stryker, 2008; Skottun et al., 1991). As a result, simple cells exhibit little response to diffuse light stimulation of the entire RF. Instead, they are optimally excited by bright (or dark) slits, bar, or edges on a background of opposite polarity (Hubel & Wiesel, 1962). Moreover, these cells are selective to specific orientations of elongated stimuli (Dräger, 1975; Hubel & Wiesel, 1959, 1962; Niell & Stryker, 2008). On the other hand, complex cell RFs could not be easily defined by two distinct ON and OFF subregions, and their responses are featured by non-linear summation of the input image (Callaway, 2001; Hubel & Wiesel, 1962; Niell & Stryker, 2008; Skottun et al., 1991). Complex cells are also optimally excited by slits, edges and bars of specific orientations (Hubel & Wiesel, 1962; Niell & Stryker, 2008). However, unlike simple cells which are

excited by a stimulus placed at a specific part of the RF, complex cells usually have larger RFs and can be excited by a stimulus at their preferred orientation placed at any part of the RF (Hubel & Wiesel, 1962). Therefore, Hubel and Wiesel (1962) have proposed that simple cell RFs were constructed by converging the circular RFs of LGN neurons that are appropriately aligned, while complex cells pool inputs from simple cells with similar orientation preferences with different signs and positions of their RFs. The LGN - simple cell circuitry has been experimentally demonstrated (Chapman et al., 1991; Reid & Alonso, 1995). whereas the simple cell - complex cell circuitry is still under debate (Callaway, 2001).

Building on the pioneering studies of Hubel and Wiesel (1959, 1962), how neurons in the visual cortex respond to a broader variety of parameters in the visual scene have been further characterised, using similar simple, parametric and two-dimensional visual stimuli such as moving bars, drifting gratings and noise stimuli (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). Following the early studies in primates and carnivores, the rodent visual cortex has also been extensively investigated. Both simple and complex cells have been identified in the mouse visual cortex (Niell & Stryker, 2008). Apart from orientation selectivity, the size and structure of the RFs as well as the ability to integrate responses over time lead to V1 neurons of rodents, carnivores and primates also selectively respond to stimuli with specific spatial frequencies, temporal frequencies, contrasts, and directions of motion (mice: Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell and Stryker, 2008; cats: DeAngelis et al., 1993; Frazor et al., 2004; Glezer et al., 1982; Pasternak et al., 1995; Zhang et al., 2017; primates: Bredfeldt and Ringach, 2002; De and Horwitz, 2022; Foster et al., 1985; Guan et al., 2021; Gur and Snodderly, 2007; Mazer et al., 2002). In the mouse visual cortex, the majority of V1 excitatory neurons show clear orientation selectivity, and inhibitory neurons are generally less tuned to orientation (Gao et al., 2010; Niell & Stryker, 2008). Simple cells usually only respond to stimuli within a specific range of spatial frequencies that typically range from 0.02 to 0.08 cpd, and the phase and orientation of the stimuli needs to align with the ON and OFF subregions of their RFs (Gao et al., 2010; Niell & Stryker, 2008). The majority of V1 neurons are tuned to temporal frequency (Gao et al., 2010). Almost half of V1 neurons show low-pass properties with the median preferred temporal frequency at 1.2Hz, and the rest are bandpass neurons with the median at 1.9 Hz (Gao et al., 2010). The spatial and temporal frequency tuning indicates the speed tuning properties of V1 neurons, as speed of visual motion can be calculated by the ratio of the temporal and spatial frequencies. Almost half of the V1 neurons are tuned to speed of moving random-dot patterns (Gao et al., 2010). V1 display an intermixed distribution of visual motion speed preferences, with neighbouring V1 neurons preferring distinct spatial and temporal frequencies spanning a diverse range (Andermann et al., 2011; Marshel et al., 2011). However, invariant speed tuning seems to be rare in V1 of mice compared to

macaques (LeDue et al., 2012; Priebe et al., 2006). In mouse V1, most neurons have a fixed preferred spatial frequency, leading to their preferred speed varying with temporal frequency (LeDue et al., 2012).

HVA neurons generally have similar properties as V1 neurons yet with more functional specialisation (Glickfeld & Olsen, 2017). Consistent with the notion of circuit hierarchy, HVA neurons usually have larger RFs than V1 neurons (DSouza et al., 2022; Murgas et al., 2019; van Beest et al., 2021; Wang & Burkhalter, 2007). The spatiotemporal tuning of HVAs show more specialisation compared to V1 (Andermann et al., 2011; Han et al., 2022; Marshel et al., 2011; Tohmi et al., 2014). For example, compared to V1, neurons in PM prefer lower temporal frequencies and higher spatial frequencies, whereas neurons in AL prefer higher temporal frequencies and lower spatial frequencies (Andermann et al., 2011; Marshel et al., 2011). LM, AM, and LI prefer both spatial and temporal frequencies to be higher (Marshel et al., 2011; Tohmi et al., 2014). As a result of their spatiotemporal preferences, HVAs such as AL and PM which show specialised biases towards fast-moving and slow-moving stimuli respectively, whereas LM, AM and LI prefer stimuli moving at intermediate speeds (Andermann et al., 2011; Marshel et al., 2011; Tohmi et al., 2014).

There are several interesting differences regarding how V1 neurons are functionally organised across species. Layer 4C of the primate and carnivore V1 is organised in alternating ocular dominance columns, where neurons each column receive LGN inputs predominantly from one of the eyes (Bonhoeffer & Grinvald, 1991; Hubel & Wiesel, 1969; Hubel & Wiesel, 1965, 1972; LeVay et al., 1975; Weliky et al., 1996). The orientation preferences of neurons in V1 are also organised in a columnar structure – neurons with the same orientation preferences reside within the same cortical column, and the orientation preference varies linearly when traversing the cortex parallelly (Hubel & Wiesel, 1962; Hubel & Wiesel, 1968, 1974; Hubel et al., 1977). Orientation columns are proposed to be organised around pinwheel centres (singularity) with weak orientation selectivity, and neurons surrounding the singularity show orientation preferences that change smoothly in a radial fashion (Bartfeld & Grinvald, 1992; Baxter & Dow, 1989; Blasdel, 1992; Braitenberg & Braitenberg, 1979; Dow, 2002; Obermayer & Blasdel, 1993). Similar to mice, V1 of primates generally follows a retinotopic organisation (Adams & Horton, 2003; Tootell et al., 1988). On the contrary, the absence of clear ocular dominance columns in the rodent visual cortex has been widely acknowledged, especially given their small binocular zone (Adams & Horton, 2009). No large-scale periodic organisation of orientation preferences were observed in rodents either (Bonin et al., 2011; Ohki et al., 2005; Van Hooser et al., 2005). Nevertheless, a recent study revealed some spatial clustering of ipsilateral and contralateral eye dominated neurons across different cortical depths in the binocular V1 of mice Ringach et al., 2016, suggesting potential columnar organisation of the rodent visual cortex which may have been overlooked before.

The properties of V1 neurons, as examined using simple two-dimensional stimuli, have led to the widely accepted view that V1 acts as a set of Gabor filters on two-dimensional

retinal images (Bonin et al., 2011; Niell & Stryker, 2008). V1 was perceived as a collection of spatiotemporal filters selective for specific parameters of the two-dimensional retinal image, such as spatial and temporal frequency, orientation and direction (Adelson & Bergen, 1985; Baker & Issa, 2005; Basole et al., 2003; Mante & Carandini, 2005). Progressive functional specialisation occurs along the hierarchy of the visual cortical pathway (from V1 to HVAs), allowing the visual system to process the visual scene from detecting simple features to perform complex visually-guided tasks such as object recognition, spatial perception and motion perception (Glickfeld & Olsen, 2017; Wang et al., 2012; Yu et al., 2022).

1.2.4 Locomotion-related modulation of neurons in the mouse visual cortex

A key feature of mouse visual cortex is its strong modulation by locomotion. The effects of locomotion on activity of neurons in the visual cortex are characterised extensively in the context of running (Ayaz et al., 2013; Dadarlat & Stryker, 2017; Eriskens et al., 2014; Keller et al., 2012; Niell & Stryker, 2010; Saleem et al., 2013). The response amplitude of excitatory neurons in V1 can be amplified substantially during running (Eriskens et al., 2014; Niell & Stryker, 2010). Niell and Stryker (2010) recorded units in upper layers of V1 when mice were running while viewing open-loop grating stimuli. They found that visually-evoked responses of most units were significantly elevated once their running speed surpassed 1 cm/s, while spontaneous neuronal activity were only increased in a small proportion of neurons. Once the mice started to run, the increase in response amplitude was not correlated with the running speed. Similar conclusions were drawn by Eriskens et al. (2014). Although they found that spontaneous activity was also elevated around locomotion onset, the enhancement in visually-driven responses was much stronger. Meanwhile, locomotion modulation did not change the selectivity towards orientation, contrast or spatial frequency of neurons (Eriskens et al., 2014; Mineault et al., 2016; Niell & Stryker, 2010). Therefore, it has been proposed that locomotion induces a modulation of cortical state which boosts the gain of sensory responses (Bennett et al., 2013; Harris & Thiele, 2011; Niell & Stryker, 2010). Running state-induced modulation in visually-evoked neuronal activity were later found in all layers of V1 (Dadarlat & Stryker, 2017; Eriskens et al., 2014), although the effect of running was most prominent in L2/3 during adulthood compared to other layers (Hoy & Niell, 2015). A higher proportion of neurons were found to be suppressed by running in deeper layers of V1 compared to superficial layers (Eriskens et al., 2014).

This widespread gain modulation across V1 can improve encoding of visual information. Compared to the stationary periods, running boosted the mutual information between neuronal responses and visual stimuli (quantified by directions and orientations of drifting gratings) in all layers of V1, and L2/3 enjoys the greatest fractional increase (Dadarlat & Stryker, 2017). On the population level, parameters of visual stimuli (grating

movement direction and grating orientation) were more accurately decoded from population activity in all layers of V1 during running (Dadarlat & Stryker, 2017). Running also reduced the noise correlations between neuronal pairs within V1, indicating a decrease in stimulus-independent trial-to-trial variability of neuronal responses and enabling a more robust neural representation of visual stimuli in V1 (Dadarlat & Stryker, 2017; Erisken et al., 2014). The neuronal activity in the population also became less correlated during running when mice were trained on a visual discrimination task to detect changing in stimulus contrast (McBride et al., 2019), which allows neurons to convey information in a more independent manner and improve perception sensitivity (Ruff & Cohen, 2014).

However, other studies suggest that running-related modulation may not be a generic gain modulation. Firstly, running can alter selectivity of neurons towards sensory stimuli. Running was found to modulate spatial integration of V1 neurons (Ayaz et al., 2013). Most V1 neurons are suppressed by larger stimuli due to surround suppression, when the sensitivity to a stimulus decreases when the stimulus is surrounded by a similar pattern (Adesnik et al., 2012; Van den Bergh et al., 2010). Running was found to significantly reduce surround suppression in head-fixed mice viewing grating patches of various sizes, which enables V1 neurons to integrate over broader areas in the visual field (Ayaz et al., 2013). Secondly, running does not impose a uniform effect on all V1 neurons. During running, the degree of gain modulation in response to grating stimuli was larger in neurons selective for high spatial frequencies compared to neurons tuned to low spatial frequencies, potentially amplifying the spatial resolution of visual perception (Mineault et al., 2016). Moreover, V1 neurons are not merely modulated by running state but also tuned to running speed (Christensen & Pillow, 2022; Keller et al., 2012; Saleem et al., 2013). There are V1 neurons that monotonically increase or decrease their activity as a function of running speed, and neurons that are non-monotonically tuned to running speed, both with and in absence of visual stimuli (Christensen & Pillow, 2022; Saleem et al., 2013). When mice were running in a VR corridor with grating and plaid patterns, V1 neuronal responses can be described by linear integration (weighted sum) between running speed and optic flow speed (Saleem et al., 2013). Therefore, Saleem et al. (2013) proposed that instead of a binary and uniform modulation of cortical state, locomotion signals provide a drive to V1 neurons. Locomotion signals can be integrated with visual signals and can alter visual preferences of V1 neurons. On the other hand, a series of studies recorded V1 responses when mice were running in a VR environment with closed-loop visual feedback interspersed by periods of visuomotor mismatch created by sudden visual flow halts (Attinger et al., 2017; Jordan & Keller, 2020; Keller et al., 2012; Leinweber et al., 2017; Muzzu & Saleem, 2021; Zmarz & Keller, 2016). They found that a small subset of L2/3 V1 neurons were responsive (or exhibited subthreshold responses) to the visual flow halts during locomotion, but not to passive viewing of visual flow halts. This observation was interpreted in the predictive coding framework where V1 neurons are signalling the prediction error between bottom-up optic flow speed input and top-down

predictions of optic flow speed based on current running speed.

Different regions of the visual system are modulated by locomotion in distinct ways. The influence of locomotion on subcortical regions seem minimal or inconsistent compared to V1. Although Niell and Stryker (2010) did not find a significant change in LGN activity during running compared to stationary periods, Erisken et al. (2014) found that locomotion onset transiently increased both spontaneous and visually-driven activity in dLGN, and the effect was much less prominent compared to the visual cortex. When imaging neuronal responses to grating stimuli in a running mice in open loop, (Roth et al., 2016) found that the activity of $\sim 20\%$ dLGN boutons in V1 can be explained by a cooperative integration between running and visual motion speed signals. On the contrary, many LP boutons' tuning to running and visual motion speed are anti-correlated, indicating that they may signal visuomotor mismatch (Roth et al., 2016). In superficial SC, neuronal activity was not consistently modulated by running and there seem no prominent trend (increase or decrease) in the activity of the whole population (Savier et al., 2019). A slight increase in activity of a proportion of cells (20%) and slight decrease in some other cells (10%) could be observed during running compared to stationary periods, and the modulation was much weaker compared to V1 (Savier et al., 2019). Among HVAs, visually-evoked activity in LM and AL were also positively correlated with running, although to a smaller degree (Christensen & Pillow, 2022). In contrast, activity in PM, RL and AM were either unchanged or suppressed by running (Christensen & Pillow, 2022). Despite negative correlation between neuronal activity and running in some HVAs, running enhances the performance of decoding drifting gratings from population activity in all areas (Christensen & Pillow, 2022).

Locomotion affects different cell types in V1 in a distinct manner. Some studies found that while the activity of somatostatin-positive (SST) interneurons were suppressed by running in darkness, vasoactive intestinal peptide-positive (VIP) neurons (a group of interneurons that mainly inhibit SST neurons Jiang et al., 2015; Pfeffer et al., 2013), increased their responses during running with or without visual stimulation (Fu et al., 2014; Jackson et al., 2016; Reimer et al., 2014). The effect on parvalbumin-positive (PV) populations is more diverse, including both positive and negative association with locomotion (Fu et al., 2014). These results led to the proposal that locomotion-related elevation of excitatory neuronal activity in V1 is facilitated by disinhibition via VIP neuron activation (Fu et al., 2014). However, Polack et al. (2013) found that both spontaneous and visually-evoked activity of PV and SST neurons was increased during running. Pakan et al. (2016) sampling a larger population of inhibitory neurons in V1 using two-photon calcium imaging found that locomotion-related modulation of SST neuronal activity depends on the behavioural context. While VIP and PV neurons increased their activity both in darkness or under visual stimulation during running, running only consistently elevated SST neuronal activity under visual stimulation (also found by Attinger et al., 2017). The modulation can also depend on visuomotor experience. Attinger et al. (2017)

trained mice to run in either closed-loop coupled VR environments (CT mice) or non-coupled environments (NT mice). Running onsets strongly drove VIP activity in CT mice but not NT mice. These results indicate that a simple disinhibition model may not capture the full circuit mechanisms for locomotion-related modulation in V1.

Movement-related modulation of V1 activity was mainly studied in the context of running for mice. However, head, eye and even facial movements have also been found to modulate V1 activity (Bouvier et al., 2020; Guitchounts et al., 2020, 2022; Meyer et al., 2018; Parker, Abe, Leonard, et al., 2022; Stringer et al., 2019). Simulated head movements induced by body rotations and real head-orienting movements were found to suppress V1 activity in the dark while exciting V1 neurons across all cortical layers in ambient light (Bouvier et al., 2020; Guitchounts et al., 2020). There are also V1 cells tuned to direction of head movements, but their head direction tuning was not correlated between the dark and light conditions (Guitchounts et al., 2020). The effects of head movements on V1 activity could be confounded by changes in retinal image during head movements due to the lack of eye tracking data in these studies. However, Parker, Abe, Leonard, et al. (2022) used miniature head-mounted cameras to track the eye and head positions of mice in real time (Meyer et al., 2018) and obtain retinal images when corrected for gaze directions. They show that the activity of many V1 neurons can be described by a multiplicative integration between visual inputs and eye and head positions, indicating encoding of gaze directions by V1 neurons. Large-scale recordings of V1 neurons and simultaneous facial movements of mice suggest that V1 activity without explicit visual stimulation can be largely explained by spontaneous activity such as running, whisking, sniffing, and other facial movements (Stringer et al., 2019). The extent of locomotion-related modulation in mouse V1 contrasts sharply with what has been observed in primates. Spontaneous activity of primates provides minimal drive to neuronal activity in primate V1, V2 and V3, and most locomotion-related modulation can be explained by changes in retinal images (Talluri et al., 2023). Therefore, compared to mice, the integration of visual and locomotion-related signals is likely to happen in later stages of visual processing, in areas such as MT (Kim et al., 2015a; Kim et al., 2016; Nadler et al., 2008, 2009).

1.3 Neural mechanisms of depth perception

1.3.1 Neural mechanisms of depth perception from binocular disparity

The neural mechanisms of depth perception from binocular disparity has been extensively studied across species. Studying binocular disparity as a depth cue stemmed from the invention of stereoscope by Charles Wheatstone in the 1830s (Wade, 2002; Wheatstone, 1949). The stereoscope allows the observer to look into two mirrors angled at 90 degrees to each other reflecting two slightly different two-dimensional images of the same

object. The difference in the images viewed between two eyes can create a sensation of depth, establishing horizontal binocular disparity as a cue for depth perception (Qian, 1997; Wade, 2002; Wheatstone, 1839; Wheatstone, 1838; Wheatstone, 1842). Wheatstone also observed the crossed and uncrossed binocular disparity for nearby and far depths relative to the focal point. Over a hundred years later, Julesz (1971) created the random-dot stereograms, where patterns of random dots with a horizontal shifts were viewed by two eyes, generating a sensation of depth from binocular disparity. In this way, the visual stimulation can be precisely controlled, and he showed that even when monocular pictorial cues and motion parallax cues were excluded, humans can perceive depth from binocular disparity alone.

Disparity tuning in cat visual cortex

Hubel and Wiesel (1962)'s milestone study in cat striate cortex (the V1 in cats) shows that the activity of most cells can be driven by visual stimuli in either eye, and thus having a receptive field in each eye. The stimuli required to stimulate these cells are similar between two receptive fields. This study raised the possibility that binocular integration could happen as early as V1. Barlow et al. (1967) then continued to record neurons in V1 of anaesthetised cats. By moving simple patterns such as a black bar, a white slit or black-white edges around the receptive fields of binocularly-activated neurons in one or two eyes, they found neurons that responded to stimuli at a similar position in the visual field of both eyes much stronger than stimuli in only one eye. Moreover, neurons in V1 were excited by different degrees of horizontal and vertical disparities between stimuli presented to the two eyes, and their preferred disparities were distributed over a horizontal range of $\sim 6.6^\circ$ and a vertical range of $\sim 2.2^\circ$. A series of studies followed (Bishop et al., 1971; Ferster, 1981; Fischer & Krüger, 1979; Nikara et al., 1968; Pettigrew et al., 1968; von der Heydt et al., 1978) to characterise the disparity-tuned neurons in cats. Later studies have also found disparity-tuned cells in area 18 of the cat's V1 as well as in HVAs such as area 21a (Ferster, 1981; Wang & Dreher, 1996). The functional significance of cat V1 in stereoscope depth perception was confirmed by the abolishment of distinguishing between two three-dimensional shapes created by random-dot stereograms following lesions of the areas 17-18 (Ptito et al., 1992). Two types of disparity-tuned cells have been identified (Ferster, 1981; Guillemot et al., 1993; Poggio & Fischer, 1977). The first type consists of binocular cells that are either excited (tuned excitatory) or inhibited (tuned inhibitory) maximally by specific degrees of disparity, when the stimuli pass over the receptive fields of the two eyes simultaneously. The second type consists of cells broadly excited by any near disparity (near cells) or far disparity (far cells) in a band-pass manner. The activity of these cells is usually dominated by one eye.

Disparity tuning in primate visual cortex

Similarly, disparity-selective neurons have been experimentally identified in monkeys (Burkhalter & Van Essen, 1986; Felleman & Van Essen, 1987; Gonzalez, Krause, et al., 1993; Gonzalez, Relova, et al., 1993; Hubel & Wiesel, 1970; Hubel & Livingstone, 1987; Maunsell & Van Essen, 1983; Poggio & Fischer, 1977; Poggio et al., 1988; Poggio & Talbot, 1981; Poggio et al., 1985). Hubel and Wiesel (1970) have first reported disparity-tuned neurons in V2 (area 18) of anaesthetised monkeys using moving slits. Some cells were found to respond maximally to a specific displacement between the visual fields of two eyes, while not responsive to monocular stimuli. Cells with similar disparity preferences were also found to group together.

Subsequently, studies have reported disparity-tuned neurons across V1, V2, V3, V3A, V4, V5/middle temporal area (MT) and inferior temporal cortex (IT) in both anaesthetised and awake monkeys, using stimuli such as moving bars, random-dot stereograms or three-dimensional shapes (Adams & Zeki, 2001; Anzai et al., 2011; Bakin et al., 2000; Cumming & Parker, 1997; Cumming & DeAngelis, 2001; DeAngelis & Newsome, 1999; Hinkle & Connor, 2002; Hubel et al., 2015; Hubel & Livingstone, 1987; Poggio & Fischer, 1977; Poggio et al., 1988; Verhoef et al., 2010). In V1 and V2, the majority (84%) of neurons are tuned to disparity (Poggio & Fischer, 1977). In V1, V2 and MT, disparity-tuned neurons predominantly reside in separate stripes from non-orientation-selective neurons (DeAngelis & Newsome, 1999; Hubel & Livingstone, 1987). Nearby neurons usually have similar disparity preference yet different direction preferences (DeAngelis & Newsome, 1999; Hubel & Livingstone, 1987; Peterhans & von der Heydt, 1993; Roe & Ts'o, 1995), allows similar depths to be detected from objects moving at various directions. Specifically characterised in MT, the organisation of disparity-tuned neurons mirrors the orientation pinwheel structures in V1, clustering into cortical columns by preferred disparity while varying their disparity preferences smoothly parallel to the cortical surface (DeAngelis & Newsome, 1999). Both types of neurons that are sharply tuned to disparity (tuned excitatory/inhibitory cells) and neurons that are broadly tuned to near or far depths (near/far cells) were found across visual areas, allowing both fine and coarse binocular vision (DeAngelis & Newsome, 1999; Poggio et al., 1988).

The dorsal and ventral stream areas differ in processing local vs. global disparity cues. Displaying anticorrelated random-dot stereogram stimuli to both eyes by inverting the contrast of the image abolishes the sensation of depth for human observers (Cogan et al., 1993, 1995). Primate V1 and dorsal stream areas such as MT and MST exhibit inverse responses to the anticorrelated stimuli (Cumming & Parker, 1997; Krug et al., 2004; Takemura et al., 2001), yet ventral stream areas such as V4 and IT are almost silent (Janssen et al., 2003; Tanabe et al., 2004). This reflects that while the dorsal stream areas perform the low-level processing of local binocular disparity signal in the visual stimuli, only the ventral stream areas perform global computation and are more closely related to the formation of stereoscopic depth perception.

Disparity tuning in mouse visual cortex

When mice were trained to jump across a gap with variable distances, optogenetic suppression of the binocular zone the V1 significantly reduced the success rate as well as the accuracy of jumping (Parker, Abe, Beatie, et al., 2022), demonstrating the critical role of V1 in depth perception from binocular cues. Other studies have characterised disparity tuning in V1 and several HVAs in the mouse visual cortex (La Chioma et al., 2019, 2020; Scholl et al., 2013). By presenting stimuli such as drifting gratings or random-dot stereograms with a phase shift between two eyes, a majority of neurons ($\sim 50 - 80\%$) in the binocular region of V1, LM and RL show strong binocular disparity-dependent excitation or suppression (La Chioma et al., 2019, 2020). Compared to cat V1 neurons, mouse V1 neurons are modulated less by disparity (Scholl et al., 2013). There is no difference in the degree of disparity modulation between simple cells and complex cells in mouse V1, whereas in cat V1, complex cells are tuned to smaller disparities compared to simple cells (Chino et al., 1994; Scholl et al., 2013).

Depth preferences from binocular disparity are related to the elevation of RFs of neurons in V1. V1 neurons with RFs in the lower half of the visual field prefer nearer depths compared to V1 neurons representing the upper visual field (La Chioma et al., 2019). There are some fine-scale spatial clustering of neurons with similar disparity preferences (La Chioma et al., 2020). Neurons within $10\ \mu\text{m}$ from each other tend to prefer similar disparities (La Chioma et al., 2020). Interestingly, HVAs display specialisation in depth preferences from disparity – compared to V1 (estimated distances encoded by neurons recorded: 3.5 - 24.6 cm) and LM (4.0 - 18.3 cm), RL neurons prefer much nearer depths (2.6 - 7.0 cm) (La Chioma et al., 2019). Apart from the specialisation in depth preferences, some HVAs also display higher-order processing of disparity signal compared to V1 (La Chioma et al., 2020). Similar to primate dorsal stream areas, neurons in mouse V1 and RL display inverse disparity tuning to anticorrelated random-dot stimuli across the two eyes, whereas analogous to the primate ventral stream areas, neurons in mouse LM reduce its disparity tuning to anticorrelated stimuli (La Chioma et al., 2019). This indicates that V1 and RL perform low-level processing of local disparity cues whereas LM is responsible for high-level processing of global disparity cues for depth perception.

Unexpectedly, the disparity tuning of neurons in V1, LM and RL are not correlated with their ocular dominance (La Chioma et al., 2020; Scholl et al., 2013). Ocular dominance is usually measured by presenting gratings to each eye separately, and neurons in V1, LM, and RL are usually strongly driven by the contralateral eye (Dräger, 1975; Gordon & Stryker, 1996; La Chioma et al., 2020; Mrsic-Flogel et al., 2007). Three modes of disparity tuning and ocular dominance were found – neurons that are classified as monocular by ocular dominance but are strongly modulated by binocular disparity, neurons that are classified as binocular by ocular dominance with no modulation by binocular disparity, and neurons that are classified as binocular by ocular dominance and show strong tuning to binocular disparity (Scholl et al., 2013). It is possible that neurons classified

as monocular by ocular dominance still receive inputs from both eyes, yet the monocular responses elicited by the non-dominant eye could not reach the spiking threshold (Scholl et al., 2013). However, with binocular integration, the neuronal responses could reach the spiking threshold and fire (Scholl et al., 2013).

1.3.2 Neural mechanisms of depth perception from motion parallax

Depth perception from motion parallax in primates

Compared to binocular disparity, the neural mechanisms of motion parallax remain much less explored. Existing studies have focused on primates. Early studies used random-dot stereograms to test depth perception from motion parallax in rhesus monkeys and humans (Cao & Schiller, 2002, 2003; Schiller et al., 2011; Zhang & Schiller, 2008). The visual stimuli they used were a plane created by random dots rotating around a vertical axis with some protrusions on the plane, while the observer remained still. Depth was simulated by the difference in speed of motion of the dots. However, as the sign of depth (near or far relative to the fixation point) from motion parallax is ambiguous without extra-retinal inputs (the direction of visual motion of objects depends on the direction of locomotion of the observer Farber and McConkie, 1979; Hayashibe, 1991; Rogers and Graham, 1979), the behavioural performance and the neuronal responses in V1 they measured could be a pure result of tuning to visual motion speed. Unlike mouse V1, spontaneous movements in face and body do not modulate visual responses of primate V1 neurons (Stringer et al., 2019; Talluri et al., 2023), which constrains the capacity of primate V1 neurons to integrate visual motion and self-motion signals.

More recent studies have addressed this issue when studying motion parallax responses in the higher visual area MT of macaque monkeys (Kim et al., 2015a; Kim et al., 2016; Nadler et al., 2008, 2009). Monkeys were presented a circular patch of random dots simulating planes at different depths while being passively translated horizontally in front of the stimuli. The monkeys made compensatory smooth eye movements to track the fixation point during the translation. Therefore, eye movement signals can be integrated with visual motion signals to compute depth from motion parallax. Pictorial cues such as shading and relative size were eliminated. This motion parallax condition has been compared to a pure retinal motion condition, where the visual motion of the stimuli during the translation was replicated while the animal remains stationary. MT neurons are highly selective for visual motion direction and speed (Britten et al., 1993; Britten, 2003; Lagae et al., 1993; Maunsell & Van Essen, 1983; Saito et al., 1986), and show binocular disparity tuning (Bradley et al., 1995; DeAngelis & Newsome, 1999; DeAngelis & Uka, 2003; Maunsell & Van Essen, 1983). Nearly two thirds of recorded MT neurons exhibit selective responses to specific depths indicated by motion parallax cues and display symmetrical tuning to simulated depth by pure retinal motion (Nadler et al., 2008, 2009).

This indicates that the depth selectivity in MT neurons could not be purely attributed to visual motion speed tuning alone, but was a result of the integration of visual motion signals and the eye movement signals that generated the visual motion (Nadler et al., 2008, 2009). Interestingly, depth tuning from motion parallax in MT are mostly monotonic, rising or declining from near to far (Nadler et al., 2008).

Kim et al. (2015a) further linked the motion parallax selectivity of MT neurons to depth perception behaviourally. Macaques were trained to report the depth sign (near or far) from random dot stimuli while being translated horizontally, and the difficulty of the task was manipulated by introducing a proportion of dots with incoherent depths from the target depth. Macaques were able to report the depth sign from the stimuli, and their performance showed a typical psychometric curve as a function of task difficulty. The responses of many individual MT neurons (35%) could predict the animal's perceptual choice of depth from motion parallax above chance. As a population, although the sensitivity of most MT neurons towards depth (measured by the depth coherence of random dots when an ideal observer could predict 82% of stimuli correctly based on neuronal responses) was lower than the sensitivity of the behaviour of the animals, a small proportion of neurons outperformed the animal. This suggests that the perceptual choice of depth can be accounted for by the responses of small populations of sensitive MT neurons. Despite ample evidence suggesting a strong correlation of MT responses with depth perception from motion parallax, inactivation experiments are needed to link MT causally with depth perception.

Depth perception from motion parallax in rodents

The neural mechanisms of depth perception from motion parallax in other species such as carnivores or rodents are also poorly understood. Rats with bilateral lesions of visual cortex failed to show depth discrimination on the visual cliff, yet the performance was unimpaired for cats with lesions of the visual cortex (Meyer et al., 1966). This suggests that the visual cortex is required for depth perception for rats but not for cats. However, this does not address whether visual cortex is required for motion parallax processing.

A number of studies examined the depth perception of rodents (rats, mice, and Mongolian gerbils) and cats through training them to jump across a distance between two raised platforms (Carey et al., 1990; Ellard et al., 1984; Goodale et al., 1990; Lashley & Russell, 1934; Olson, 1980; Parker, Abe, Beatie, et al., 2022; Russell, 1932). Mongolian gerbils and mice were able to jump across a gap with only monocular information, showing that they can perceive depth with monocular visual cues alone (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022). Interestingly, it was observed in early studies that locusts make lateral head movements to perform visual scanning ("peering") before jumping towards a target, which has been proposed as a way of generating motion parallax cues from self-motion to assist more accurate depth perception (Sobel, 1990b; Wallace, 1959). Similarly, Mongolian gerbils and mice were also observed to make vertical translational

head movements (head bobs) before jumping across the gap (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022). The number and frequency of head bobs were positively correlated with the gap distance and the accuracy of jumping, and monocular mice performed more frequent head bobs compared to binocular mice. These results indicate that in absence of binocular cues, motion-based cues can be important for depth perception.

The head bob movement provides a behavioural readout for depth perception from motion parallax. Lesions in the visual cortex introduced severe deficits in jumping accuracy and latency to gerbils and a significant increase in the proportion of trials where head movements were performed before a jump (Ellard et al., 1986). Gerbils with SC lesions show no impairment in jumping accuracy (Ellard et al., 1986). In contrast, gerbils with lesions in the visual cortex or mice with optogenetic suppression of V1 monocular zone displayed significant disruptions of the jumping performance (Ellard et al., 1986; Parker, Abe, Beatie, et al., 2022). Mice were more likely to abort the jumping attempt (Parker, Abe, Beatie, et al., 2022). These lesion and inactivation experiments indicate that V1 is essential for depth perception in mice using monocular visual cues, whereas the role of SC in depth perception of mice is unclear. However, multiple other monocular cues such as looming, shading and perspectives still exist in the jumping task, so it is not conclusive whether V1 specifically supports motion parallax processing.

Extra-retinal signals required for depth perception from motion parallax

What extra-retinal locomotion signals are required for depth perception from motion parallax? In humans and primates, eye movement signals contribute to the disambiguation of depth sign from motion parallax (Nadler et al., 2009; Naji & Freeman, 2004; Nawrot, 2003). Depth perception of human subjects was tested using stimuli with four rows of vertical bars alternately moving in opposite directions, and the adaptation to the movements of bars would cause an illusions of stationary bars moving to the opposite direction due to motion after-effect (Nawrot, 2003). A perception of depth could be created when the subject translated their head or eyes – the bars with an illusory movement opposite to the head or eye translation appeared to stand out in depth. When the subjects translated their head while making compensatory eye movements or just translating their eyes, the perceived depth was same as expected. However, when the subjects only translated their head while keeping their eyes stationary in the head, a smooth eye movement command signal in the direction of head movement would be generated to suppress the compensatory eye movement signal, causing the subjects to perceive the opposite depth signs from the stimuli (Nawrot, 2003). The perception of depth sign was completely ambiguous when the smooth eye movement was removed (Nawrot, 2003). These experiments show that smooth eye movement command signal was required to produce a perception of depth from motion parallax.

The role of eye movement in motion parallax processing was also investigated in macaque monkeys (Nadler et al., 2009). Using a patch of random dots simulating plans

at different depths, Nadler et al. (2009) tested the neuronal responses in macaque MT with four conditions: the motion parallax (MP) condition where animals were translated while maintaining fixation by compensatory eye movements, the head only (HO) condition where only head translation was performed without eye movements, the eye only (EO) condition where only smooth pursuit eye movements were performed to follow the moving stimuli while the animal remained stationary, and the retinal motion (RM) condition where only the stimuli moved while the head and eyes of the animals remained stationary. The activity of MT neurons were tuned to depth signs in both the MP condition and the EO condition, whereas MT neurons were not selective to depth signs in the RM and HO conditions. This indicates that the depth selectivity of MT neurons from motion parallax in macaques relies on smooth pursuit eye movements as the extra-retinal signals but not the head/body translation.

Voluntary smooth pursuit eye movements have not been observed in rodents (Ambrad Giovannetti & Rancz, 2024). Instead, the eye movements of mice were mainly driven by head rotations to stabilise the direction of gaze during locomotion Holmgren et al., 2021; Meyer et al., 2020; Michael et al., 2020. Head tilt (pitch and roll) elicits compensatory non-conjugated eye movements to stabilise the visual field, whereas head yaw triggers a "saccade and fixate" gaze pattern where eye movements shift and then stabilise the gaze (Meyer et al., 2020). Therefore, self-generated visual motion in mice is primarily driven by translation of the head of the animals.

1.3.3 Neural mechanisms of depth perception from pictorial cues

There are limited studies that investigate the neural mechanism of depth perception from pictorial cues in isolation. It has been proposed that the ventral stream visual areas in humans and non-human primates are particularly key to processing monocular pictorial cues (Heider, 2000; Mon-Williams et al., 2001). For example, lesions in the occipito-temporal lobe in humans can cause visual agnosia, when the basic visual functions such as the detection of brightness, colour and motion remain intact whereas high-level visual functions such as object recognition and face recognition are disrupted (Heider, 2000; Marotta et al., 1997; Mon-Williams et al., 2001; Turnbull et al., 2004). They mainly rely on binocular cues for depth perception (Mon-Williams et al., 2001), and cannot make accurate depth judgements to perform visually-guided reaching with only monocular pictorial cues (Marotta et al., 1997) such as relative size (Mon-Williams et al., 2001). Visual agnosia patients were unable to correctly imagine three-dimensional rotations of objects, and this could be attributed to their deficiency in extracting three-dimensional structures from pictorial cues in the two-dimensional pictures of objects such as shading and linear perspective (Turnbull et al., 2004). In macaques, a large proportion of neurons (67%) in the area inferior temporal (IT) showed selectivity to the degree of tilt of surface created by either texture gradients or disparity gradients, and half of these neurons were significantly tuned to both gradients defined by texture and by disparity (Liu et al., 2004).

The tilt tuning remained invariant to texture patterns, indicating that these neurons are tuned to depth gradients rather than local two-dimensional visual patterns (Liu et al., 2004). Depth estimation from pictorial cues has not been systematically examined in the mouse visual cortex. However, we would expect that similar to primates, calculating depth from pictorial cues is likely to be supported by HVAs, as they can pool the outputs from V1 neurons that are tuned to local visual features such as spatiotemporal frequency, orientation, size and luminance (Adesnik et al., 2012; Andermann et al., 2011; Bonin et al., 2011; Dipoppa et al., 2018; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008) and calculate depth gradient by comparing population activity that represent a larger area in the visual field.

1.4 Project overview

Although depth computation is essential for animal behaviour in the three-dimensional environment and motion parallax is an important depth cue for many animals, there is a lack of systematic characterisation of the neural mechanisms underlying depth perception from motion parallax, especially in rodents. Therefore in this thesis, I aim to explore the developmental origin of depth perception and then investigate the neural mechanisms underlying depth estimation from motion parallax in mice.

In Chapter 3, I will demonstrate whether depth perception is innate for mice. I used a customised version of the visual cliff as a paradigm to systematically characterise the depth discrimination behaviour in mice. I will then show whether depth perception requires prior visual experience by testing mice born and reared in the dark on the visual cliff. I will also present a pilot experiment exploring whether superior colliculus, an important input structure to the visual cortex, is required for the computation of depth before the visual information reaches the visual cortex.

In Chapter 4, I will explore the neural mechanisms underlying depth estimation from motion parallax in visual cortex. Specifically in this project, I used V1 as a starting point, due to its well-characterised tuning properties and the extensive evidence of visuomotor integration. I focus my experiments on L2/3 of V1 due to the limitation on the range of cortical depth imageable by two-photon microscopy (Yildirim et al., 2019). I will start by presenting the design of a virtual reality (VR) environment which allows precise manipulation of visual stimuli to present motion parallax cues during active navigation of mice. Then, I demonstrate how neurons in L2/3 of V1 exhibit depth-selective responses when head-fixed mice were navigating in the VR environment using two-photon calcium imaging to record neuronal activity.

Next, in Chapter 5, I will illustrate how depth selectivity in V1 neurons can be generated from motion parallax signals. By modelling neuronal activity as a function of running speed and optic flow speed, I will show that depth selectivity arises from the integration of running speed and optic flow speed tuning.

Finally, in Chapter 6, I will characterise the representation of depth across V1. I will demonstrate that the activity of V1 neurons during active navigation can be characterised by three-dimensional RFs, in the dimensions of two-dimensional visual field as well as depth. Then, I will characterise the representation of depth across V1 and illustrate how depth is mapped by specific populations of V1 neurons.

Overall, this thesis provides novel insights on how V1 is functionally organised to drive behaviour in our environment. Not only is depth perception an evolutionarily important behaviour, depth is also encoded explicitly in the visual system to facilitate this essential computation. It shows that mouse V1 does not merely reflect the two-dimensional visual field, but represent three-dimensional locations of visual cues. It also provides a functional explanation for the widespread locomotion-related modulation of the mouse visual cortex.

Chapter 2

Methods

2.1 Authorship declaration

This chapter built upon materials that I have previously disseminated in a preprint on bioRxiv (He et al., 2024).

2.2 Animal welfare compliance

All experimental procedures were performed in accordance to the UK Animals (Scientific Procedures) Act of 1986 (PPL PP4882546) and approved by the Animal Welfare Ethical Review Body at the Francis Crick Institute. Animals used in each experiment are detailed below.

2.3 Examining the innate nature of depth perception using the visual cliff test

2.3.1 Visual cliff

The visual cliff assay was conducted on a custom-built arena adapted based on previous studies (Fox, 1965; Gibson & Walk, 1960; Trychin & Walk, 1964). Each mouse was placed on a raised central platform (raised by 6 cm and 5 cm wide) and would need to make a choice and descend to either the shallow side or the deep side. The two sides were created by a large transparent panel divided by the central platform (25 cm $\times$ 25 cm on each side) made of acrylic and were raised 60 cm from the floor. On one side, a checkerboard texture (25 cm $\times$ 25 cm), printed on paper and affixed to an acrylic board, was placed directly beneath the panel, creating the "shallow side". On the opposite side, another checkerboard texture (25 cm $\times$ 25 cm) was placed 60 cm below (on the floor of the chamber), simulating a vertical depth on the "deep side". The whole setup were supported by two metal rods and bases. All acrylic boards were cut with laser to achieve their precise sizes and were joint by UV-curing transparent glue.

Textures of various densities were used across different experiments. For the experiments with textures of identical physical density, the checkerboard textures placed on both the shallow and deep sides consisted of $2.5\text{ cm} \times 2.5\text{ cm}$ squares (Figure 3.1A-B). 16 C57BL/6J (wild-type, JAX stock #000664) mice were used. For the experiments with textures of matched density from the perspective of the animal, the textures placed on the shallow side consisted of $2\text{ cm} \times 2\text{ cm}$ squares, and the textures on the deep side consisted of $22\text{ cm} \times 22\text{ cm}$ squared (Figure 3.2A-C). 8 C57BL/6J mice were used. For the experiments to confirm the effect of texture density on the animals decision on the visual cliff, a texture consisting of $1\text{ cm} \times 1\text{ cm}$ squares and a texture consisting of $2\text{ cm} \times 2\text{ cm}$ squares were placed on the two sides respectively, and both sides were made to be shallow (Figure 3.3A). 8 C57BL/6J mice were used.

To limit the difference in illumination and shadows between the two sides, a string of dimmable LED lights were arranged in arrays to cover the whole back panel of the setup, and light diffuser sheet was placed in front of the light bulbs to create diffused lighting for the whole chamber (Figure 3.1B). The left and right interior panels of the chamber were also covered by diffuser sheets to eliminate pictorial cues on the chamber walls (Figure 3.1B).

At least 1 hour before the start of the visual cliff experiments, mice were briefly restrained and their whiskers were trimmed with a shaver to prevent them from making the decision using tactile cues. Mice were tested on the visual cliff for 1 session without any prior exposure to the visual cliff setup. Each session consisted of either 12 or 20 trials, and had equal number of trials with the shallow side (or the deep side) appearing on the left and right sides of the setup. The trials where the shallow side appearing on the left or right side of the setup was interleaved in a pseudorandom order. Before each trial, the texture indicating the shallow side was placed on the left or right side of the setup according to the predetermined sequence, while the texture indicating the deep side remained on the floor. At the beginning of each trial, a mouse was placed on the central platform and was allowed to freely navigate on the arena. The choice of side for the mouse was determined when mice descended to one side of the acrylic platform with all four paws. Once the mouse had made its decision, it was removed from the arena and the top surfaces where the mouse had made contact with were cleaned with Anistel wipes and tissues to eliminate the remaining smell. Mice were allowed a maximum of 3 attempts to complete a trial. For the very first trial, each mouse was given up to 3 attempts until they made a decision, each lasting 10 minutes. If no decision was made after 3 attempts, the trial was recorded as "no descent". In subsequent trials, each mouse was allowed up to 3 attempts, each lasting 5 minutes, before being counted as "no descent" if a decision was not made. A session was terminated if a mouse failed to make a decision for 3 consecutive trials. A new trial was initiated as soon as a decision had been made or once a "no descent" trial had been confirmed.

2.3.2 Dark rearing

3 pregnant C57BL/6J female mice were transferred to dark cages a week before their due date, while 2 pregnant female C57BL/6J mice were transferred to normal cages in the same room. 23 pups of the dams in dark cages were born and raised in dark cages for 10 weeks with no exposure to light (dark-reared mice), whereas 17 pups of the dams in normal cages were born and raised under normal lighting condition for the same duration (control mice). Dark-reared mice were fed and checked daily by the experimenter with infrared night vision goggles.

During the visual cliff experiments, the experimenter was blinded on whether the tested animal was from the dark-reared or the control cohort. At the start of each visual cliff test, one mouse was transferred into a different cage by another experimenter, and was placed in the experimental chamber for 10 minutes to habituate to the lighting level. Then the visual cliff assay would be operated in the same manner indicated in Section 2.3.1.

2.3.3 Visual cliff with optogenetic inhibition

To test whether the activity of superior colliculus is required for depth perception in mice, 5 Vgat-cre mice (JAX stock #028862, Vong et al., 2011) were injected with AAV-hSyn-DIO-Chrmine-P2A-mScarlet (titre: 7.97×10^{12} vg/ml) and 5 other Vgat-cre mice acting as controls were injected with AAV-FLEX-tdTomato (titre: 1.2×10^{13} vg/ml) bilaterally in their SC. Each side had two injections of 124 nl of virus in total, with the two anterior injection sites at ± 1.15 mm in the medial-lateral axis from lambda, 0.5 mm anterior to lambda, and 2 mm from the brain surface, and the two posterior injection sites at ± 1.15 mm in the medial-lateral axis from lambda, 0.1 mm anterior to lambda and 1.8 mm from the brain surface. Then, a Dual Fiber-Optic Cannula containing two optic fibres (TFC_300/330-0.37_2mm_TS2.3_FLT Two Ferrule Cannula, Doric, core diameter 200 μm) was implanted so that the ends of the two optic fibres targeted the brain at ± 1.27 mm in the medial-lateral axis from lambda, 0.35 mm anterior to lambda and 1.12 mm below the brain surface. The optic fibre implants allowed photo-stimulation of Vgat-positive cells. The location of the fibres implantation was confirmed via postmortem fluorescence imaging, and the ends of the fibres were located in the centre of the superior colliculus on the superficial layer. Based on the size and location of the viral injections, the target area of viral infection was estimated to be $\sim 500 \mu\text{m}^2$ on each side.

A red laser (Doric Laser Diode Fiber Light Source, $\lambda = 638\text{nm}$, maximum power 80mW) was used to deliver the optogenetic stimulation. The laser was connected with a rotary joint (Doric Fiberoptic Rotary Joint, Intensity Division) using a Mono Fiberoptic Patchcord (MFP_200/240/LWMJ-0.22_1m_FCA-FC, Doric). Before a behavioural trial, the laser and the rotary joint was connected to the optic fibre implant on the head of the mouse with another Mono Fiberoptic Patchcord (MFP_200/240/900-0.22_1m_FC-

ZF1.25, Doric).

20 trials were performed on the visual cliff for each mouse as described Section 2.3.1. During half of the trials, the laser was turned on ("laser-on" trials) as soon as the mouse was placed on the central platform. The laser was triggered with pulses of square waves at 20 Hz with a 5% or 15% duty cycle. The maximum average laser power was 9.8 mW measured at the end of the Mono Fiberoptic Patchcord connecting to the optic fibre implant with $\sim 85\%$ coupling efficiency (from the Patchcord to the Cannula). During the other half of the trials, the laser was turned off ("laser-off" trials). The laser-on and laser-off trials were interleaved in a pseudorandom order.

Since the sample size for this experiment was quite small (4 mice for each group), to control for animal-to-animal variability, I used hierarchical bootstrap (Saravanan et al., 2020) to test the statistical significance of data from this experiment. For each bootstrap sample I resampled mice followed by trials for each mouse. Then I calculated the statistic of interest for each bootstrap sample. When determining whether mice at a specific condition significantly preferred the shallow side, I calculated the average probability of choosing the shallow side across all sampled trials, and then subtract the chance level (0.5) from the mean value. When comparing whether the performance of mice at two conditions (e.g. optogenetic vs. control mice with the laser on), I calculated the difference in the average probability of choosing the shallow side across all sampled trials between the two conditions. Finally, I determined the quantile q of the bootstrap distribution at 0 and calculated the two-sided bootstrap p-value as $2 \min\{q, 1 - q\}$.

2.4 Characterising the neural representation of depth in the primary visual cortex

2.4.1 Animals

Experiments described by all figures apart from Figure 4.6 were conducted in 7 transgenic mice, including 6 $\text{Emx1-Cre} \times \text{Ai95D}$ mice (JAX stock #028865, Madisen et al., 2015 and #005628 Gorski et al., 2002) and one $\text{CamKII-tTA} \times \text{TRE-GCaMP6s}$ line G6s2 mouse (JAX stock #003010, Mayford et al., 1996 and #024742 Wekselblatt et al., 2016) expressing genetically encoded calcium indicators GCaMP6f or GCaMP6s respectively in cortical excitatory neurons. Both male and female mice were used (1 male and 6 females). In Figure 4.6, 2 C57BL/6J (JAX stock #000664) female mice expressing GCaMP8m in their primary visual cortex through virus injections were included.

2.4.2 Cranial window implantation

To implant the cranial window for chronic two-photon calcium imaging, mice were anaesthetized with isofluorane (1.5-2.5%). A subcutaneous injection of dexamethasone (0.02 ml at 2 mg/ml) was administered to reduce brain oedema. The skull was exposed,

and a metal headplate was affixed with dental cement. A craniotomy (4 mm in diameter) was performed over the left visual cortex, and the removed bone flap was replaced by a 4 mm glass coverslip. Imaging was typically conducted when the mice were 3 to 7 months (up to 10 months).

2.4.3 Virus injection

For the two mice expressing GCaMP8m in their primary visual cortex through virus injections, a craniotomy (4 mm in diameter) was first performed over the left visual cortex as described above. Then, virus injections were performed through a glass micropipette mounted to a Nanoject (Nanoject III, Drummond Scientific). The glass pipette was pulled to a fine tip (outer diameter $\sim 10\text{-}20\text{ }\mu\text{m}$) prior to the surgery, and was backfilled with an adequate amount of the viral mixture (1:5 AAV1-Syn-GCaMP8m, 1:10 AAV2/1-Ef1a-FLEX-H2B mCherry and 1:7000 AAV2/1hSyn-Cre). The micropipette was mounted onto the Nanoject attached to a stereotax prior to the surgery. After the craniotomy was performed, the micropipette was positioned at -2.8 medial/lateral and -3.5 anterior/posterior from bregma. Then, the micropipette was lowered into the cortex at $300\text{ }\mu\text{m}$ from the dura surface. 100 nl of the viral mixture was injected at a speed of 1 nl/s interspaced with 3 s pauses. 2 minutes after a round of viral injection was completed, the micropipette was further lowered to $550\text{ }\mu\text{m}$ from the dura surface, and another 100 nl of the viral mixture was injected following the same procedure. After two rounds of viral injections were completed, the micropipette was withdrawn from the brain, and a 4 mm glass coverslip was placed over the left visual cortex to replace the removed bone flap.

2.4.4 Virtual reality setup and visual stimuli

During the recordings, mice were head-fixed on a polystyrene cylindrical wheel and presented with a virtual reality (VR) setup on four monitor screens (MSI Optix G241) surrounding the mouse. The monitors were arranged in portrait orientation approximately along four sides of a hexagon centred on the mouse and covered a visual field of ~ 240 degrees horizontally and ~ 80 degrees vertically. The animals' position in the VR environment was updated in closed loop based on the distance they traveled on the wheel recorded using a rotary encoder (Kubler 2400). The virtual reality environment was rendered in Bonsai software (Lopes et al., 2015) using the BonVision package (Lopes et al., 2021). The 3D environment was rendered using cube mapping, which projected a 360° visual scene onto the view ports around the mouse (Lopes et al., 2021). The view ports were defined as the positions of the screens relative to the mouse, whose translation and rotational vectors were calibrated using ArUco markers (Bradski, 2000; Garrido-Jurado et al., 2014) displayed on the screens and placed at the location of the mouse. The right eye was monitored with a camera (Basler acA1440-220um) synchronized with the two-photon acquisition ($\sim 15\text{ Hz}$) and illuminated with an infrared light placed above the camera.

The VR environment had a gray background (11.9 cd/m^2), a 5 cm wide floor with a checkerboard texture positioned 2.5 cm below the mouse in the VR, and 24 black spheres presented at varying virtual distances from the animal. The locations of the centres of the spheres were selected in cylindrical coordinates (Figure 4.1F). The longitudinal axis of the cylinder was aligned to the animal's path of travel in the VR, while its radius was equal to the virtual depth of a given trial chosen pseudorandomly from a range of logarithmically spaced values (either 5 virtual depth values between 6 cm and 600 cm, or 8 depths between 5 and 640 cm). The spheres were spaced along the axis of the animal's locomotion by $0.15 \times d$ with 19 spheres ahead of and 5 spheres behind the mouse. The cylindrical azimuth of the spheres, which determines their position relative to the horizon, was randomly chosen from a uniform distribution between -40 and 40 degrees or between 140 and 220 degrees, i.e. around the horizon on either the left-hand side or right-hand side of the animal (Figure 4.1F). To maintain constant visual stimulation during the trial, whenever the animal moved the distance of $0.15 \times d$ in the VR, resulting in a sphere disappearing outside of the monitors' field of view behind the mouse, a new sphere was generated $2.7 \times d$ ahead of the mouse at a randomized cylindrical azimuth. New spheres gradually faded in from gray over 0.3 s. The virtual radius of the spheres on trials of different virtual depths was equal to $\tan 5^\circ \times d$, so that they covered the same angular extent in the visual field of the mice across different virtual depths (10 degrees when the mice passed closest to each sphere in the VR). The spheres were a solid black color without specular highlights.

Each trial continued until the mouse traveled 6 m after which the spheres faded to gray and a 10 s inter-trial period commenced. A probabilistic reward of soy milk ($\sim 8 \mu\text{l}$ of 10% SMA Wysoy) was delivered from a spout in front of the mouse 2 s into the inter-trial interval on 60-80% of trials. Each session consisted of at least 10 trials for each virtual depth.

During open loop sessions, the animal's movement on the wheel was recorded but did not control its virtual position in VR. Instead, the running trajectory of a previous closed loop session of the animal was replayed and determined the animal's moment-to-moment virtual position and the updating of the sphere stimulus.

The nominal refresh rate of the four monitors was 144 Hz. To synchronize the visual stimuli with imaging data, and measure the actual frame rate, a small square ($3 \text{ cm} \times 3 \text{ cm}$) with varying luminance was displayed in the bottom left corner of the leftmost monitor. The luminance changes of this square were recorded with a photodiode at 1 kHz (HARP Behavior device, Champalimaud Research Scientific Hardware Platform). For 2 mice (25 sessions), the square alternated between black and white at every frame and was used to detect dropped frames. To determine the closed loop latency between rotary encoder inputs and updates of the visual stimulus, the brightness of the square was updated following an irregular predefined sequence of 5 values for the remaining 5 mice (60 sessions, Figure 4.2A). The sequence was designed to alternate brightness increases

and decreases in consecutive frames. Frames were then detected based on the photodiode trace. To identify which frame was presented at each time point, the cross-correlation between the filtered predefined sequence and the photodiode recording was computed to continuously measure display lag.

2.4.5 Calibration of screen positions

The positions of the four screens were calibrated using ArUco markers as described in previous literature Bradski, 2000; Garrido-Jurado et al., 2014. First, the intrinsics of the camera used for calibrating the screen positions were calculated by taking pictures of a flat checkerboard pattern (consisting of $3.1 \text{ cm} \times 3.1 \text{ cm}$ squared) at various angles (Figure 4.3A, Kshirsagar et al., 2022). The camera was modelled as a pinhole camera with distortions, projecting 3D coordinates of real-world objects (e.g. the corners of the checkerboard) to two-dimensional coordinates on the image. The coordinates of the corners of the checkerboard patterns were detected on the images. The undistorted normalized coordinate of a corner on a image (x, y) is a result of applying the translation (T) and rotation matrices (R) of the checkerboard pattern relative to the camera as well as the intrinsics of the camera (K) to the predefined 3D coordinates of the corners of the checkerboard (X, Y, Z) :

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K \begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (2.4.1)$$

where s is a scaling factor. The camera intrinsics matrix (K) include parameters such as focal lengths (f_x, f_y) and optical centers (c_x, c_y) in pixels:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}. \quad (2.4.2)$$

The actual coordinates of checkerboard corners would incorporate distortion factors including radial distortion (parts of the image bulging outward or squeezing inward relative to the optical center) and tangential distortion (asymmetrical shifts on the image due to misalignment of lens to the image plane). Therefore, the camera intrinsics can be estimated by iteratively optimising the pinhole camera project from the real-world 3D geometry of the checkerboard pattern to the known two-dimensional coordinates of checkerboard corners across multiple pictures, while minimising reprojection error. Next, two ArUco markers with known patterns and dimensions ($5 \text{ cm} \times 5 \text{ cm}$) were placed in the middle of each monitor screen and the location where the head of mice would be fixed, respectively (Figure 4.3B). A picture containing both markers was taken. The corners and x , y , and z axes of the individual markers were detected from the picture using the

OpenCV package (Figure 4.3B, Bradski, 2000), and the translation (T_i, T_j) and rotation vectors (R_i, R_j) of the aruco markers relative to the camera were computed. Then, the relative translation (R_{rel}) and rotation (T_{rel}) between the two aruco markers could be computed as:

$$R_{rel} = R_j^T R_i, \quad (2.4.3)$$

$$T_{rel} = R_j^T (T_i - T_j), \quad (2.4.4)$$

which indicates the transformation of screens from the perspective of the mice.

2.4.6 Eye camera calibration

Four cameras were used to monitor the eye, facial, and body movements of the mice. To monitor the position of the right eye of the mice, the position of the camera monitoring the right eye was calibrated. First, the intrinsics of the camera were estimated by taking pictures of small checkerboard patterns (consisting of $0.8 \text{ mm} \times 0.8 \text{ mm}$ squares) using the camera from various angles, and calculated using the method described in Section 2.4.5 (Figure 4.3C). The relative rotation from the eye to the camera was estimated by taking a picture of an ArUco marker ($1 \text{ cm} \times 1 \text{ cm}$, Figure 4.3D), with the method described in Section 2.4.5. As the changes in visual motion speed resulting from gaze direction shifts were the primary focus, the translation vector of the eye camera was not relevant in this context.

2.4.7 Behavioral training

To encourage mice to run in the VR environment, they underwent a 2-4 week training period under food restriction. Mice were acclimatized to the behavioral setup for at least 3 sessions of 5-15 min. Mice initiated their training with a minimum distance traveled per trial of 1 m that was gradually increased to 6 m depending on the animals' running behavior. Each training session lasted 30 min - 1 h. Imaging commenced once the mice could reliably complete ~ 2 trials per minute. After the training period, food restriction continued during the imaging experiments.

2.4.8 Two-photon calcium imaging

Fluorescence signals from neurons in layer 2/3 of the primary visual cortex ($150\text{-}300 \mu\text{m}$ below pia) were recorded using a custom-built resonant scanning two-photon microscope (Independent Neuroscience Services) and a Nikon 16x water-immersion objective (NA 0.8). The microscope was controlled with ScanImage 2020 software. GCaMP6f and GCaMP6s fluorescence signals were excited with a 930 nm laser (SpectraPhysics MaiTai DeepSee eHP) at 70-100 mW and acquired with a 525/50 emission filter (FF03-525/50-25, Semrock). Imaging frames (1024×1024 pixels) were acquired at ~ 15 Hz. Individual

recordings covered a field of view of $\sim 661 \mu\text{m} \times 661 \mu\text{m}$. To minimize the interference from the monitors' during imaging, a custom electronic circuit activated the backlights of the monitors (screen blanking) only during the turn-around period of the resonant X mirror.

2.4.9 Widefield calcium imaging

The locations of V1 and higher visual areas (HVAs) were determined using wide-field calcium imaging of the whole visual cortical window. Widefield calcium imaging was performed using a camera integrated into the two-photon microscope (Independent NeuroScience Services) using a Nikon AF Nikkor 85mm f/1.8D camera lens as the objective and a 100 mm tube lens (Thorlabs AC508-100-A-ML). Two excitation LEDs at 470 nm (M470L3, Thorlabs, with excitation filter FF02-447/60-25, Semrock) and 405 nm (M405L4, Thorlabs, with excitation filter FF01-405/10-25, Semrock) were combined using a dichroic mirror (FF458-Di02-25x36, Semrock) and coupled to the infinity focused imaging path (FF495-Di03 BrightLine primary dichroic mirror). Images were acquired with an emission filter (FF03-525/50-25, Semrock) and a CMOS camera (Basler acA2040-120um) at 60 Hz. Illumination of the two LEDs was interleaved for hemodynamic correction using a microcontroller (Arduino Uno) triggered by the camera's FlashWindow output.

During imaging, mice were head-fixed on the wheel and presented with a circular square-wave grating patch of 20 degrees diameter (at a spatial frequency of 0.2 cpd and a temporal frequency of 2 Hz) at 30 or 90 degrees azimuth and 0 degrees elevation on a black background. Each grating patch was presented for 4 seconds and grating direction (8 directions in 45 degree increments) changed every 500 ms in pseudorandom order. The grating patches were repeated 50 times for each location in pseudorandom order with 8 seconds of blank screen between presentations.

Images were first deinterleaved to separate 405 nm and 470 nm channels. Hemodynamic signal correction was carried out by performing linear regression on each pixel of the 470 nm channel's signal using the 405 nm channel signal as the independent variable. The residuals of the regression added to the mean fitted signal of the 470 nm channel was used as the corrected widefield signal for each pixel. Corrected signals were detrended by applying a high-pass filter of 0.001 Hz. The signal at each pixel was normalized by its average value. The responses to the grating patches at each retinotopic location were averaged across repetitions to determine the locations of V1 and HVAs.

2.4.10 Grating stimuli

To map the spatiotemporal tuning of V1 neurons, I recorded the neuronal responses to 2D grating stimuli in a subset of sessions ($N = 6$ sessions from 2 mice). Sinusoidal drifting gratings were presented on a black background (0.12 cd/m^2) rendered in Bonsai software (Lopes et al., 2015) using the BonVision package (Lopes et al., 2021). Full-field drifting

gratings were displayed exclusively on the two screens covering the right visual field of the mice. The stimuli consisted of gratings with combinations of 6 spatial frequencies (0.01, 0.02, 0.04, 0.08, 0.16, 0.32 cpd) and 6 temporal frequencies (0.5, 1, 2, 4, 8, 16 Hz) drifting at 8 directions (0, 45, 90, 135, 180, 225, 270, 315, 360 degrees). The gratings were presented in a pseudorandom order, with each combination of spatial frequency, temporal frequency, and direction displayed for 4 seconds before switching to the next. 0° represented vertical gratings moving forward. Every session included 5 repetitions of all grating combinations.

2.4.11 Data analysis

To account for the nested structure of the experimental data, statistical comparisons were conducted using hierarchical bootstrap (Saravanan et al., 2020) unless specified otherwise. In analyses at the level of individual neurons, for each bootstrap sample I resampled mice, followed by experimental sessions for each mouse, and neurons for each session and calculated the statistic of interest, such as correlation coefficient between two variables, or difference in medians between two conditions, for each sample. I then determined the quantile q of the bootstrap distribution at 0 and calculated the two-sided bootstrap p-value as $2 \min\{q, 1 - q\}$.

Eye movements

To determine if the visual stimulation triggered eye movements, in a subset of sessions ($N = 16$ sessions from 2 mice) I reconstructed the gaze direction during behaviour. Eyelid position, the reflection of the infrared light and 12 key points around the pupil border were tracked on each frame using DeepLabCut (Mathis et al., 2018). Frames in which the eye was closed (distance between top eyelid and bottom eyelid $<$ mean distance - 3 standard deviations, or DeepLabCut likelihood < 0.88) were excluded. For the remaining frames, the 12 markers around the pupil were fitted with an ellipse which was used to perform gaze reconstruction as previously described (Wallace et al., 2013). Briefly, the eye was modeled as a 3D sphere with a circular pupil rotating at a constant radius around the the eye centre. Translation of the eye relative to the camera was corrected using the infrared light reflection as the origin for each frame. The position of the eye centre on the camera frame was defined as the least square estimate of the intersection of all minor axes of the fitted ellipses. The scale factor, which depends on the radius of the eye, was estimated based on the ratio of the short and long axis of the fitted ellipses and the position of the eye centre (Wallace et al., 2013). With these parameters, the pupil radius r and rotation angles relative to the camera axes, ϕ and θ define the reprojected pupil border in the camera frame. I performed a grid search for each frame to find the (r, ϕ, θ) minimizing the difference between the fitted and reprojected ellipses.

To estimate the azimuth and elevation from the gaze vector in camera coordinates, I calibrated the camera position using an ArUco marker (Garrido-Jurado et al., 2014)

placed parallel to the floor in the camera field of view and pointing in the animal's direction of travel. The pose of the marker relative to the camera was estimated using OpenCV (Bradski, 2000) and used to rotate the gaze vector from camera coordinates to world coordinates.

Angular velocity (Figure 4.8K) was calculated on the mean filtered azimuth and elevation traces (rolling window of 3 frames). Saccades (Figure 4.8L) were defined as instances when the filtered velocity traces (calculated on median-filtered azimuth and elevation traces with a rolling window of 5 frames) exceeded 75 degrees per second, i.e. a 5 degrees change in a frame.

Preprocessing of two-photon imaging data

Imaging frames were registered and segmented using the suite2p package (Pachitariu et al., 2017) with anatomical segmentation (Cellpose Stringer et al., 2021) to avoid bias towards active cells. Fluorescence traces were detrended by subtracting a rolling baseline calculated as the 20th percentile in a moving window of 900 frames. To remove neuropil contamination of the fluorescence trace for the ROIs, I fitted the fluorescence of the ROIs and their surrounding areas with asymmetric Student-t distributions. The mean of these distributions depended on a shared neuropil signal that affected both ROI and surrounding fluorescence (Orsolic et al., 2021). Finally, I estimated F_0 , by fitting a Gaussian mixture model with 2 components to the fluorescence time series and identified F_0 as the mean of the smaller Gaussian. $\Delta F / F_0$ was defined as $\Delta F / F_0 = (F - F_0) / F_0$, where F was the fluorescence value after neuropil correction.

Depth selectivity

To determine the preferred virtual depth for each ROI, I calculated the mean $\Delta F / F_0$ across each trial for each neuron, and fitted a Gaussian model to the average single trial responses f and the corresponding log-transformed virtual depth d displayed in each trial:

$$f = ae^{\frac{-(d-d_0)^2}{2\sigma^2}} + b, \quad (2.4.5)$$

where a was the peak response amplitude, d_0 was the log-preferred depth, σ was the tuning width, and b was the baseline fluorescence. The preferred virtual depth d_0 was constrained between $\ln 2$ and $\ln 2000$ cm, and the peak amplitude a was constrained to be positive. The tuning width σ^2 was constrained to >0.5 to avoid overfitting. All depth tuning curves were plotted using this fit.

To identify depth-selective neurons, I fitted the Gaussian model using 5-fold cross validation. On each fold, 80% of trials were assigned to the training set, which was used to estimate model parameters and 20% were assigned to the test set, which was used to evaluate model predictions. Then I calculated the Spearman's correlation coefficient of predicted and observed fluorescence on test trials across all 5 folds. Depth-selective

neurons were defined as having a Spearman's correlation p-value < 0.05 and a Spearman's correlation coefficient > 0.1 .

Near-preferring cells and far-preferring cells (Figure 4.5D) were defined as the cells with preferred virtual depths close to the bounds of d_0 (log-preferred virtual depth $< \ln 2 + 10^{-4}$ for near-preferring cells, and $> \ln 2000 - 10^{-4}$ for far-preferring cells).

Optic flow and running speed tuning

To visualise the optic flow speed and running speed tuning curves in Figure 5.1B,F, I first smoothed the sum of responses from all frames and the number of frames at different bins of running speed or log-transformed optic flow speed respectively, using a 1D Gaussian kernel with a standard deviation of 1 bin width (10 cm/s for running speed, ~ 0.691 for log-transformed optic flow speed). The tuning curves were calculated by dividing the smoothed sum of response by the smoothed number of frames to account for the number of samples in different bins. Only frames with a running speed above 1 cm/s were included for the optic flow speed tuning curves to facilitate the logarithmic transformation of the optic flow speeds.

To determine how optic flow and locomotion-related signals were integrated and estimate preferred optic flow and running speeds, I fitted the $\Delta F/F_0$ trace as a function of log-transformed optic flow speed v and running speed r on individual imaging frames when mice were moving at > 1 cm/s using 5 models.

The optic flow only and running speed only models were Gaussian functions of log-optic flow speed v and log-running speed r , respectively:

$$\Delta F/F_0 = ae^{-\frac{(v-v_0)^2}{2\sigma_v^2}} + b, \quad (2.4.6)$$

$$\Delta F/F_0 = ae^{-\frac{(r-r_0)^2}{2\sigma_r^2}} + b. \quad (2.4.7)$$

The additive model fitted responses as the sum of Gaussian functions of log-optic flow speed v and log-running speed r :

$$\Delta F/F_0 = a_v e^{-\frac{(v-v_0)^2}{2\sigma_v^2}} + a_r e^{-\frac{(r-r_0)^2}{2\sigma_r^2}} + b. \quad (2.4.8)$$

The conjunctive model consisted of a bivariate Gaussian tuned to both log-optic flow speed v and log-running speed r :

$$\Delta F/F_0 = ae^{-\left(k_v(v-v_0)^2 + k_r(r-r_0)^2 + k_{vr}(v-v_0)(r-r_0)\right)} + b, \quad (2.4.9)$$

where

$$k_v = \frac{\cos^2 \theta}{2\sigma_1^2} + \frac{\sin^2 \theta}{2\sigma_2^2}, \quad (2.4.10)$$

$$k_r = \frac{\sin^2 \theta}{2\sigma_1^2} + \frac{\cos^2 \theta}{2\sigma_2^2}, \quad (2.4.11)$$

$$k_{vr} = \frac{\sin 2\theta}{2\sigma_1^2} - \frac{\sin 2\theta}{2\sigma_2^2}. \quad (2.4.12)$$

a was the peak amplitude of the overall response, b was the baseline activity, and v_0 and r_0 were the log-preferred optic flow and log-preferred running speeds. For the conjunctive model, θ determined the orientation of the axes of the bivariate Gaussian, and σ_1 and σ_2 were the standard deviation for its two axes.

The ratio model was a Gaussian function of the difference between log-running speed and log-optic flow speed, equivalent to the logarithm of their ratio:

$$\Delta F/F_0 = ae^{\frac{-(r-v-d_0)^2}{2\sigma_d^2}} + b, \quad (2.4.13)$$

where d_0 was the log-preferred virtual depth and σ_d was the depth tuning width. This model is similar to the depth tuning model in Eq. 4.4.1 fitted on individual imaging frames rather than trial-averages.

The preferred optic flow speed e^{v_0} was constrained between 0.03 to 3000 degrees/s. The preferred running speed e^{r_0} was constrained between 0.5 to 500 cm/s. d_0 was constrained between 0.0095 and 9.5×10^5 cm according to the lower and upper bounds of e^{v_0} and e^{r_0} . σ_v^2 , σ_r^2 , σ_1^2 , and σ_2^2 were constrained to be >0.25 . θ was constrained between 0 and 90 degrees. a , a_r , and a_v were constrained to be >0 .

To determine which model best described neuronal activity for individual neurons, I fitted each model with 5-fold cross validation. On each fold, individual trials were assigned to the training set (80%), which was used to estimate model parameters, or test set (20%), which was used to evaluate model predictions. I then computed the predicted fluorescence $\widehat{\Delta F/F_0}$ on test trials using parameters estimated on the training set for each fold and compared it to observed $\Delta F/F_0$ to compute R^2 as the fraction of variance explained by the model:

$$R^2 = 1 - \frac{\text{Var}[\Delta F/F_0 - \widehat{\Delta F/F_0}]}{\text{Var}[\Delta F/F_0]}. \quad (2.4.14)$$

The best model for each neuron was selected as the model with the highest R^2 . To compare performance between models across the dataset, I first calculated the proportion of neurons best fitted by each model for each recording session and then used hierarchical bootstrap (Saravanan et al., 2020) to resample mice and sessions 20000 times. I then compared the difference of the median proportions between models.

To estimate the preferred running speed and optic flow speed for each neuron, I fitted the $\Delta F/F_0$ trace, the corresponding optic flow speed and running speed from all trials using the conjunctive model. Closed loop recordings and open loop recordings were fitted separately. The neurons plotted in Figure 5.4A-C and Figure 5.5A-B were the depth-selective neurons well-fitted by the conjunctive model in closed loop recordings (cross-validated $R^2 > 0.02$ computed as described above, 5,737 of 31,013 depth-selective neurons recorded in closed loop conditions from 84 out of 85 sessions; no neurons passed

this threshold in 1 session). To quantify the correlation between preferred depth, preferred running speed, preferred optic flow speed or the ratio between preferred running and optic flow speed in Figure 5.4A-C and Figure 5.5A-B, I first calculated the Spearman's correlation coefficient between the two variables and calculated the p-value using hierarchical bootstrap by resampling mice, sessions, and neurons as described above.

To compare the preferred running speed and preferred optic flow speed of depth-selective neurons in closed loop vs. open loop trials in Figure 5.7E-F, I chose depth-selective neurons with good fits to the conjunctive model (cross-validated $R^2 > 0.02$) in both closed and open loop conditions (1,234 of 10,850 depth-selective neurons recorded in open loop from 34 sessions). Then, I computed the Spearman's correlation coefficient between the variables and calculated the p-value using hierarchical bootstrap by resampling mice, sessions, and neurons as described above.

To compare the peak response amplitudes of depth-selective neurons in closed loop vs. open loop trials in Figure 5.7G-I, I calculated the peak response amplitude using the conjunctive model fit (sum of amplitude a and offset b). The selection criteria for neurons was the same as Figure 5.7E-F. Then I compared the median ratio between the peak response amplitude of closed loop and open loop trials to 1 using hierarchical bootstrap by resampling mice, sessions, and neurons as described above.

Spatiotemporal and direction tuning in response to 2D grating stimuli

To determine the preferred speed tuning of V1 neurons mapped by 2D drifting gratings, I calculated the average response of individual neurons during each 4-s presentation of gratings with specific combinations of spatial frequency, temporal frequency and direction. Then, the trial-averaged neuronal responses were then fitted with a model that combined a spatiotemporal tuning component represented by a bivariate Gaussian function and a direction tuning component based on a von Mises distribution. The bivariate Gaussian component, $G(s, t)$, modelled the neuron's spatiotemporal tuning as:

$$G(s, t) = ae^{-(k_s(s-s_0)^2 + k_t(t-t_0)^2 + k_{st}(s-s_0)(t-t_0))} + b, \quad (2.4.15)$$

where

$$k_s = \frac{\cos^2 \theta}{2\sigma_1^2} + \frac{\sin^2 \theta}{2\sigma_2^2}, \quad (2.4.16)$$

$$k_t = \frac{\sin^2 \theta}{2\sigma_1^2} + \frac{\cos^2 \theta}{2\sigma_2^2}, \quad (2.4.17)$$

$$k_{st} = \frac{\sin 2\theta}{2\sigma_1^2} - \frac{\sin 2\theta}{2\sigma_2^2}. \quad (2.4.18)$$

s and t were the log-transformed spatial and temporal frequency of the presented grating. a was the peak response amplitude of the spatiotemporal tuning component, b was the baseline, and s_0 and t_0 were the log-preferred spatial frequency and temporal frequency. θ determined the orientation of the axes of the bivariate Gaussian, and σ_1 and σ_2 were

the standard deviation for its two axes. The direction tuning component, $D(\alpha)$, modelled the neuron's direction tuning as:

$$D(\alpha) = \frac{e^{(\kappa \cos(\alpha - \alpha_0))} + (1 - dsi) \cdot e^{(\kappa \cos(\alpha - \alpha_0 - \pi))}}{e^{(\kappa)} + (1 - dsi) \cdot e^{(-\kappa)}},$$

where α was the direction of the presented grating. α_0 was the preferred grating direction, and κ controlled the width of the direction tuning. dsi was the direction selectivity index ranging from 0 to 1, defined as $\frac{D_{preferred} - D_{null}}{D_{preferred} + D_{null}}$, where $D_{preferred}$ represented the preferred direction of motion and D_{null} represented the opposite direction. The trial-averaged neuronal responses was modelled as $G(s, t) \cdot D(\alpha)$ with an offset representing the baseline activity.

Receptive field estimation

To estimate the receptive field for each neuron, I used the recorded sphere locations and the animal's position in the VR to reconstruct the three-dimensional visual stimulus presented for each imaging frame. The first two dimensions corresponded to retinotopic location in spherical coordinates with a resolution of 5 degrees in azimuth and elevation, while the third dimension corresponded to the virtual depth of the stimulus on a given trial. I then constructed a design matrix representing the visual stimuli S , where each row contained the stimulus reconstruction for the corresponding imaging frame and for each virtual depth, as well as a column of 1s to account for a bias term. The number of rows corresponded to the number of imaging frames. Therefore, the dimension of S was $N_{frames} \times (N_{pixels} * N_{depths}) + 1$, where N_{pixels} was the number of pixels on each visual stimulus frame, N_{depths} was the number of virtual depths, and the N_{frames} was the number of imaging frames. S was reconstructed based on the stimuli presented on either the two screens on the right-hand side of the mouse or the two screens on the left-hand side of the mouse. I used a linear model to fit $\Delta F/F_0$:

$$\Delta F/F_0 = S\mathbf{b} \tag{2.4.19}$$

I used the regularized pseudoinverse method (Cossell et al., 2015; Smyth et al., 2003) to impose a smoothness constraint on the receptive field in azimuth and elevation and in virtual depth. The constraints aimed to make the Laplacian of the receptive field close to zero at all points. To impose this constraint, I constructed matrices L_{xy} and L_{depth} . L_{xy} aimed to smooth the receptive fields across azimuth and elevation. To assemble the matrix L_{xy} , I first generated Laplace matrices that contained the values

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \tag{2.4.20}$$

corresponding to adjacent x and y pixels for a single depth embedded in $N_y \times N_x \times N_{depth}$ matrix of zeros for each x and y pixel location and depth. These matrices were flattened and assigned to individual rows of L_{xy} .

L_{depth} controlled the smoothness the receptive fields across different virtual depths. To assemble L_{depth} , I generated Laplace matrices that contained the values $\begin{bmatrix} -1 & 2 & -1 \end{bmatrix}$ corresponding to adjacent depths for a single pixel embedded in a $N_y \times N_x \times N_{depth}$ matrix of zeros for each x and y pixel location and depth. These matrices were flattened and assigned to individual rows of L_{depth} .

I appended a column of 0s to L_{xy} and L_{depth} so as not to regularize the bias term. I then assembled the design matrix X containing the stimulus matrix S and the smoothness constraints L_{xy} and L_{depth} :

$$X = \begin{bmatrix} S \\ \lambda_{xy} \cdot L_{xy} \\ \lambda_{depth} \cdot L_{depth} \end{bmatrix}. \quad (2.4.21)$$

λ_{xy} and λ_{depth} are scalars that control the strength of regularization of receptive field smoothness across azimuth and elevation (λ_{xy}) and across virtual depths (λ_{depth}). I constructed an augmented response vector by appending 0s to the $\Delta F/F_0$ fluorescence vector corresponding to each row of L_{xy} and L_{depth}

$$\mathbf{y} = \begin{bmatrix} \Delta F/F_0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \quad (2.4.22)$$

To estimate the receptive field I found the least squares solution for the equation $\mathbf{y} = X\mathbf{b}$ as

$$\hat{\mathbf{b}} = (X^T X)^{-1} X^T \mathbf{y}. \quad (2.4.23)$$

The first N_{pixels} elements of $\hat{\mathbf{b}}$ correspond to coefficients of the receptive field, while the last element represents the bias term.

To optimize hyperparameters λ_{xy} and λ_{depth} , I used 5-fold cross validation to split imaging trials into training and test sets and estimated $\hat{\mathbf{b}}$ on each training set. I then computed the predicted fluorescence

$$\hat{\mathbf{y}} = S_{test} \hat{\mathbf{b}} \quad (2.4.24)$$

for each test set and found the combination of λ_{xy} and λ_{depth} which yielded the highest fraction of variance explained by the model for each ROI by searching over a 13 x 13 grid of values logarithmically spaced between 2.5 and 10240. The receptive field for each neuron was calculated based on the best combination of λ_{xy} and λ_{depth} by averaging the

coefficients $\hat{\mathbf{b}}$ across 5 folds. The peak azimuth and elevation of the receptive field (RF_{azi} and RF_{ele}) were defined as the position with the maximum coefficient value.

To identify the neurons with significant receptive fields, I used the procedure above to optimize λ_{xy} and λ_{depth} and fit the receptive field using the stimuli in the right visual field (RF_{right}). To estimate the null distribution for receptive field coefficients for each neuron, I used these values of λ_{xy} and λ_{depth} to fit the receptive field using stimuli in the left visual field (RF_{left}), which have the same statistics but are not expected to drive neuronal activity recorded in the left visual cortex. I calculated the mean and standard deviation of the coefficients in RF_{left} and defined the ROIs with significant receptive fields as those with a maximum value of the RF_{right} which was 6 standard deviations above the mean of RF_{left} . The 6 standard deviations threshold was chosen such that $\sim 5\%$ of neurons passed this threshold if it was applied to RF_{left} .

Cortical location

To determine the spatial location of imaged neurons within V1 in individual animals, I first found the location of neurons within the field of view (FOV) for each recording session based on their cell masks. I then aligned the location of each FOV within the whole visual cortical window by matching the blood vessel patterns between the surface of the FOVs and the overview image of the window acquired using the widefield camera.

To align the location of imaged neurons across animals, I assumed that the RF azimuth and elevation followed the same linear gradient in V1 across animals. Under this assumption, the overview location $x_{overview}$ and $y_{overview}$ of individual neurons depends linearly on their RF azimuth and elevation with a translational offset that can vary across mice:

$$x_{overview} = b_{1,1}RF_{azi} + b_{1,2}RF_{ele} + b_{1,mouse} \quad (2.4.25)$$

$$y_{overview} = b_{2,1}RF_{azi} + b_{2,2}RF_{ele} + b_{2,mouse} \quad (2.4.26)$$

I used Huber regression to fit the coefficients and individual mouse offsets ensuring robustness to outliers. I then chose a reference mouse and aligned the neurons' coordinates of all other mice to the reference mouse by subtracting the offset of the original mouse and adding back the offset of the reference mouse:

$$x_{aligned} = x_{overview} - b_{1,mouse} + b_{1,reference} \quad (2.4.27)$$

$$y_{aligned} = y_{overview} - b_{2,mouse} + b_{2,reference} \quad (2.4.28)$$

Depth selectivity in relation to RF location

To illustrate the spatial distribution of RF azimuth, elevation and the preferred depth in Figure 6.4B-D, I plotted these parameters of all depth-selective neurons with a significant RF at the location of individual neurons. In the lower panels, I smoothed the

spatial distribution map with a $113 \mu\text{m}$ Gaussian kernel. To set the transparency of the smoothed map, I calculated the sum of Gaussian weights for all plotted neurons for each pixel, normalized by its maximum value. The alpha for each pixel was set to either 5 times this value or 1, whichever was smaller.

As the virtual distance between the spheres and the mouse varied based on the spheres' positions, in the analyses in Figure 6.5D-E I calculated the corrected log-preferred virtual depth d'_0 , corresponding to the distance to the spheres at the centre of the neurons' receptive field at the peak of their depth tuning curve (Eq. 4.4.1):

$$e^{d'_0} = \frac{e^{d_0}}{\sqrt{\sin^2 RF_{azi} \cos^2 RF_{ele} + \sin^2 RF_{ele}}}. \quad (2.4.29)$$

To test for the presence of a gradient in preferred depth as a function of azimuth and elevation, I used hierarchical bootstrap (Saravanan et al., 2020) to resample neurons across mice and recording sessions. For each bootstrap sample, I performed linear regression between the RF_{azi} and RF_{ele} of the centre of the neurons' receptive fields and the neurons' corrected log-preferred virtual depth:

$$d'_0 = k_{azi} RF_{azi} + k_{ele} RF_{ele} + b, \quad (2.4.30)$$

where k_{azi} and k_{ele} defined the gradient of depth as a function of azimuth and elevation of the neurons' receptive fields, respectively.

To test for statistical significance, I modeled the coefficients (k_{azi}, k_{ele}) obtained from different bootstrap samples using a bivariate Normal distribution and used a multivariate generalization of the Z-test to compare their mean to 0. To this end, I first computed the mean $\boldsymbol{\mu}$ and covariance Σ of bootstrap samples (k_{azi}, k_{ele}). As the square of the Mahalanobis distance $\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}$ follows the chi-squared distribution with 2 degrees of freedom, the p-value is calculated as $e^{-\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}/2}$.

Population decoding

To decode virtual depth from population activity, I trained a linear SVM classifier using $\Delta F/F_0$ values of simultaneously recorded neurons in a session. I split imaging trials into training (64%), validation (16%) and test (20%) sets with 5-fold cross validation such that each trial was included in one test set. I then trained the classifier on individual imaging frames from trials in the training set using the validation set to optimize the hyperparameter C . I used the test sets to evaluate the performance of the optimized classifier to obtain the accuracy and the confusion matrix of predicted and true virtual depths. Confusion matrices in Figures 4.7A-D and 5.8A-D show decoder performance as the proportion of imaging frames for each true depth.

To determine how decoding accuracy depends on the animals' running speed, I trained a single decoder using data across all running speeds for each session. I then evaluated

decoder performance on imaging frames from trials in the test set belonging to different running speed bins. Figure 4.7E shows mean decoding error across sessions. Confidence intervals were computed using bootstrap by resampling sessions.

To compare decoding accuracy between closed loop and open loop trials, I trained and evaluated separate SVM classifiers on closed loop and open loop data.

Chapter 3

Depth perception in mice does not require prior visual experience

3.1 Authorship declaration

I, Yiran He, performed the majority of the experiments and analysis in this chapter. Marcelo Moglie from Iacaruso Lab at the Francis Crick Institute performed the surgical procedures for optogenetic mice. Xavier Cano-Ferrer (a member of the Cricks Making Lab) contributed to the design and construction of the visual cliff setup.

3.2 Background

Depth perception is a critical ability supporting the everyday life of animals in our three-dimensional environment. Whether this vital visual function is an innate ability across the animal kingdom has drawn considerable attention. The innate nature of depth perception has been widely studied using the visual cliff (Birch et al., 1982; Fox et al., 1980; Gibson & Walk, 1960; Tsuruhara et al., 2014; Walk & Gibson, 1961). This behavioural paradigm assesses depth discrimination by observing animals' tendency to avoid a visual drop. The apparatus typically features a large transparent panel divided by a central platform. On one side, known as the "shallow side," a texture is placed directly beneath the panel. On the opposite "deep side," another texture is positioned several feet below, creating a sensation of vertical depth. Animals with normal depth perception would choose to descend to the shallow side and avoid the deep side, possibly due to a fear of heights (Gibson & Walk, 1960; Walk & Gibson, 1961). The visual cliff experiment is an ideal behavioural paradigm to probe whether depth perception is innate or requires prior visual experience. It does not require prior training or high-level motor coordination for animals to perform the task (unlike the jumping tasks used in Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022). Depth discrimination can be easily quantified by animals' choices between two simple options, and the behavioural performance is robust across many species (Gibson & Walk, 1960; Walk & Gibson, 1961).

On the visual cliff, human infants display a fear of vertical drops as early as 4 months old (Birch et al., 1982; Fox et al., 1980; Gibson & Walk, 1960; Tsuruhara et al., 2014). Many birds and mammals, including chicken, sheep, goats, and rats, demonstrate an innate ability to perceive depth on the visual cliff (Cheney & Crow, 1969; Gibson & Walk, 1960; Walk et al., 1957; Walk et al., 1965). Immediately after birth or following rearing in a completely dark environment, these animals inherently preferred to descend to the shallow side, indicating that their depth perception does not depend on prior visual experience. On the other hand, some carnivores and primates seem to need visual experience to fully develop their depth perception. Compared to kittens that could actively explore the environment and receive closed loop visual feedback, kittens that only received passive visual inputs did not develop paw placing reflexes and did not avoid apparent vertical drops (Hein et al., 1970). Infant monkeys reared in the dark for more than a month could not discriminate between the shallow and deep sides on the visual cliff (Fantz, 1965).

Previous work has shown that adult mice with normal vision can reliably discriminate depth on the visual cliff, whilst mice with retinal degeneration lost this ability (Fox, 1965). Adult mice can also perform a modified version of the visual cliff task with mainly binocular information, where mice climbed down a pole and mainly engaged their upper visual field (Boone et al., 2021). However, it is unknown whether depth perception is an innate ability for mice. In addition, previous studies have used patterns with the same physical texture density for both deep and shallow sides. This means that mice could discriminate depth by comparing relative size or density of textures between the two sides, rather than relying on motion parallax or binocular disparity cues. The effect of key pictorial cues, such as relative size, on depth discrimination of mice on the visual cliff has not been systematically examined.

In this chapter, I will explore whether depth perception is innate or requires any prior visual experience for mice using the visual cliff. I will use a modified version of the visual cliff which enables flexibility in manipulating the density and height of the textures presented on both sides. If depth perception is proven innate in mice, the mouse visual system can serve as an ideal model to study the functions and connectivity of specialised neuronal populations defined by their gene expression patterns.

Another question I will touch on in this chapter is which parts of the visual system are necessary for depth perception. In the introduction, I have presented extensive anatomical and functional evidence that the visual cortex is a likely neural substrate for depth perception in mice. The visual cortex, especially the primary visual cortex (V1), is indispensable for rodents to perform depth perception tasks such as the visual cliff and jumping across a horizontal gap (Ellard et al., 1986; Meyer, 1963; Parker, Abe, Beatie, et al., 2022). The neural mechanism of depth perception from binocular disparity has been well-characterised in mice, and many neurons in V1, AL and RL display selectivity towards disparity of visual stimuli presented to two eyes (La Chioma et al., 2019, 2020; Scholl et al., 2013). The visual cortex also receives rich locomotion-related signals

from areas such as the thalamus, the hippocampal formation and secondary motor cortex (Froudarakis et al., 2019; Leinweber et al., 2017; Saleem et al., 2018; Vélez-Fort et al., 2018) and neuronal activity in the visual cortex is highly modulated by locomotion (Keller et al., 2012; Niell & Stryker, 2010; Saleem et al., 2013). This could allow neurons in the visual cortex to compute depth from motion-based depth cues by integrating visual and locomotion-related signals. However, the innate nature of depth perception across many species suggests that it may be supported by evolutionarily conserved regions such as subcortical areas. It remains unclear which subcortical structures are required for depth perception.

The two main subcortical structures that provide visual information to the visual cortex are the lateral geniculate nucleus (dLGN) and superior colliculus (SC), which are the key subcortical centres that support the image-forming pathway and the non-image-forming pathway, respectively (Juavinett et al., 2020; Seabrook et al., 2017). dLGN provides key inputs of retinal visual information to V1 (Petros et al., 2008). Given that depth computation requires perception of the visual environment and that V1 is required for depth perception in rodents, it is likely that dLGN is involved in depth perception. SC is an evolutionarily-conserved, integrative sensorimotor structure that receives inputs from the majority of RCGs. It channels the alternative pathway for retinal information to reach V1 via its projections to dLGN and higher-order thalamic nuclei (Bickford et al., 2015; Gale & Murphy, 2014; Roth et al., 2016; Seabrook et al., 2017; Tohmi et al., 2014). SC is a key structure involved in visuomotor transformations and visually guided orienting, controlling eye movements, head movements, escape and hunting behaviour (Hoy et al., 2019; Ito & Feldheim, 2018; Shang et al., 2015, 2019; Sparks et al., 1990; Wei et al., 2015). However, lesion studies suggest that superior colliculus is not required for depth perception in rodents including Mongolian gerbils and rats (Cheney & Crow, 1969; Ellard et al., 1986). Lesions in the superior colliculus of Mongolian gerbils did not significantly affect depth judgements during a task of jumping across a gap, while lesions in the visual cortex impaired the task performance (Ellard et al., 1986). Gerbils with superior colliculus lesions displayed similar latency before jumping as the control group, although their landing positions were less accurate (Ellard et al., 1986). Lesions in the superior colliculus of rats did not affect their performance on the visual cliff (Cheney & Crow, 1969). It is unknown whether SC is dispensable for depth perception in mice. Therefore, I will present a pilot experiment using the visual cliff to examine the role of SC in depth perception.

3.3 Mice can discriminate depth using visual cues on the customised visual cliff setup

I first designed and built a customised visual cliff setup (Figure 3.1A-B) based on previous studies (Fox, 1965; Gibson & Walk, 1960; Trychin & Walk, 1964). A mouse was placed on a raised central platform and could choose to descend to either a shallow side or a deep side. The two sides were created by a large transparent panel divided by the central platform. On one side, a checkerboard texture was placed directly beneath the panel, creating the "shallow side"; whereas on the opposite side, another texture was positioned 60 cm below, simulating a vertical depth on the "deep side". The textures could be easily replaced by sliding from the side of the large transparent panel, enabling us to alter texture pattern or density as well as randomising the side of texture presentation.

I aimed to establish a baseline for the performance of mice with normal vision on the visual cliff. I tested 8 adult C57BL/6J mice (referred as wild-type mice below) on the visual cliff with checkerboard textures of the same physical density on both the shallow side and deep side (Figure 3.1A-B). Mice were tested without any prior exposure to the visual cliff for only 1 session consisting of 12 trials. Before each trial, the position of the shallow and deep sides was shuffled randomly between the left and right side of the arena, while ensuring that the shallow side appeared on the left and right sides an equal number of trials. Each trial began with a mouse being placed on the raised central platform and ended when the mouse made a decision to descend to either the shallow or deep side, determined by all four paws landing on one of the sides. On average, wild-type mice with normal vision chose to descend to the shallow side in 84.4% of the trials, significantly above chance level (Figure 3.1C,E, one sample signed-rank test, $p = 0.0078$). This indicates a strong preference to the shallow side in mice while visual cues are present.

Non-visual cues, such as the smell left by the exploration of mice in previous trials, could bias their choice on the visual cliff. To examine the effect of any non-visual cues on the performance of mice, I tested another 8 wild-type mice on the visual cliff in the dark. The dark environment had no visible lighting but only infrared red light for monitoring with cameras. The frequency of choosing the shallow side was significantly lower than that in the normal lighting condition (Figure 3.1D,E, rank sum test, $p = 0.0005$), and was not different from chance level (Figure 3.1E, one sample signed-rank test, $p = 0.25$). This shows that mice rely on visual cues to discriminate depth on the visual cliff.

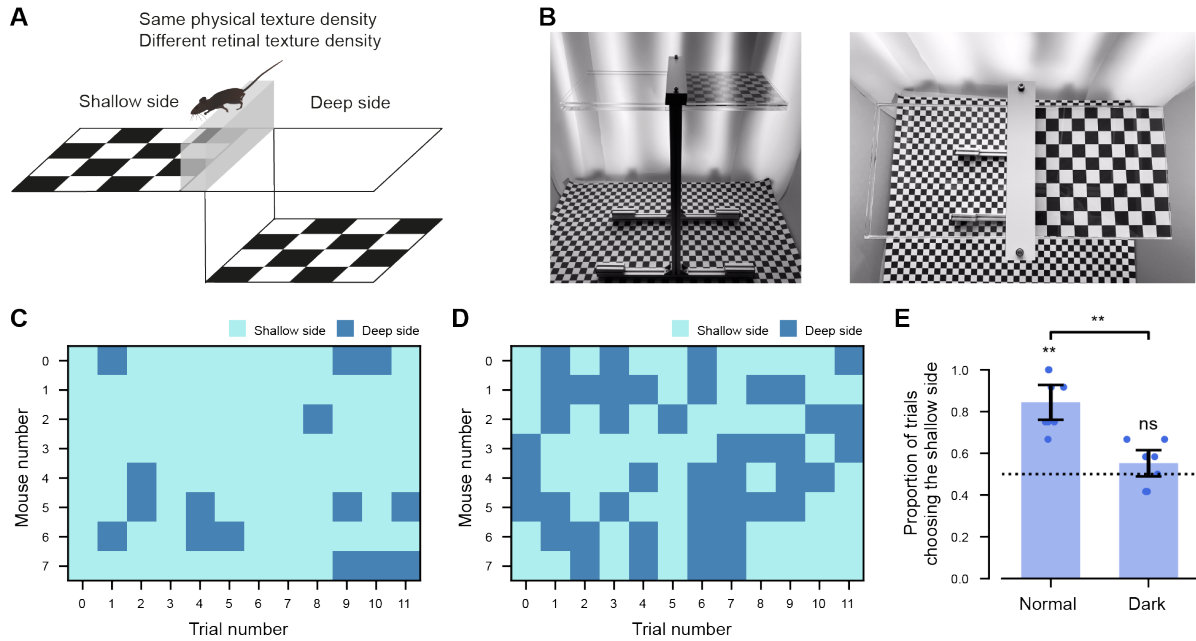


Figure 3.1: Mice can discriminate depth using visual cues on the customised visual cliff setup. (A) Schematic of the visual cliff setup with textures of the same physical density on both sides. (B) Visual cliff setup, front view (left) and aerial view (right). (C) Side of descending for each trial for each mouse on the visual cliff with normal lighting. (D) Side of descending for each trial for each mouse on the visual cliff in dark environment. (E) Proportion of trials choosing the shallow side (mean $\pm$ 95% confidence interval) in 12 trials for mice tested in normal lighting condition or dark environment ($N = 8$ for each group). **: $p < 0.01$, ns: $p > 0.05$.

3.4 Mice can discriminate depth on the visual cliff with limited pictorial cues

With the same physical texture density at both side, the texture at the deep side will appear to be smaller than the shallow side from the perspective of the mouse. Therefore, I next aimed to limit pictorial depth cues on the visual cliff test to prevent mice from simply discriminating depth based on relative size. To this end, I scaled the texture density of the checkerboard patterns on the shallow and deep side so that each square on the checkerboard covered the same visual angle from the perspective of the mice (Figure 3.2A). I also noticed that the metal base of the visual cliff setup was visible from the deep side but not the shallow side, which may bias the choices of the mice (Figure 3.1B). Therefore, I modified the setup such as the bottom texture covered the entirety of the floor, hiding the metal base of the setup (Figure 3.2B-C).

I then tested another cohort of 8 wild-type mice on the visual cliff with matched texture density in longer sessions (20 trials). Across the whole session, mice still exhibited a strong preference to the shallow side, with 90.6% of trials descending to the shallow side (Figure 3.2D-E). The probability of choosing the shallow side did not differ between the first 12 trials and the whole 20 trials (first 12 trials: 90.6%; 20 trials: 90.6%). The probability of descending to the shallow side when texture density was matched from the perspective of

3.4. Mice can discriminate depth on the visual cliff with limited pictorial cues

the mouse was similar to the condition when texture density were not matched (Figure 3.1E, 3.2E, Wilcoxon rank sum test, $p = 0.357$). These results suggest that mice are able to discriminate depth on the visual cliff when pictorial cues were limited.

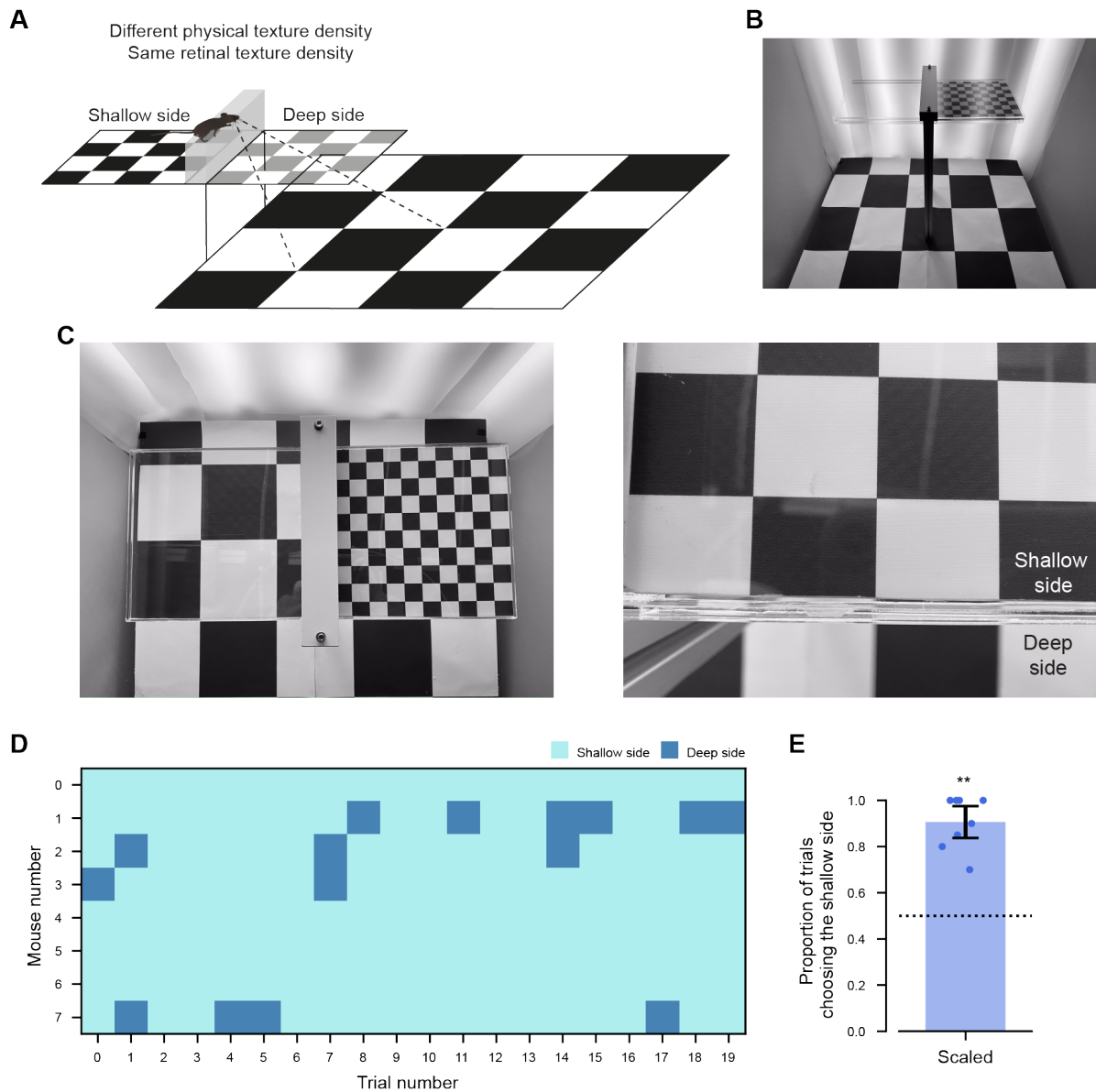


Figure 3.2: Mice can discriminate depth on the visual cliff with limited pictorial cues. (A) Schematic of the visual cliff setup with matched retinal texture density. (B) Front view of the visual cliff setup with matched retinal texture density. (C) Aerial view (left) and from the eye level of mice (right) of the visual cliff setup in (B). (D) Side of descending for each trial for each mouse on the visual cliff with matched retinal texture density. (E) Proportion of trials choosing the shallow side (mean $\pm$ 95% confidence interval) in 20 trials ($N = 8$ for each group). **: $p < 0.01$, ns: $p > 0.05$.

3.5 Texture density does not bias the decision of mice on the visual cliff

In the previous section, I show that texture density is not required for mice to discriminate depth on the visual cliff. Since the texture density will change slightly as mice move their head around during the freely-moving task, I then asked if texture density alone is sufficient for mice to judge depth on the visual cliff. I tested another 8 wild-type mice on the visual cliff with texture of different densities placed on the shallow platform of both sides (Figure 3.3A). Neither the performance of mice on the first trial (Figure 3.3C, one sample signed-rank test, first trial vs. chance level: $p = 0.641$) nor the average across all 20 trials showed significant preference on either side (Figure 3.3D, one sample signed-rank test, average across trials vs. chance level for mice completing all 20 trials: $p = 0.0656$; for all mice: $p=0.397$). This confirms that texture density alone is not sufficient to bias the decision of mice on the visual cliff, and the slight change in texture density during free exploration on the visual cliff is unlikely to bias the choice of mice.

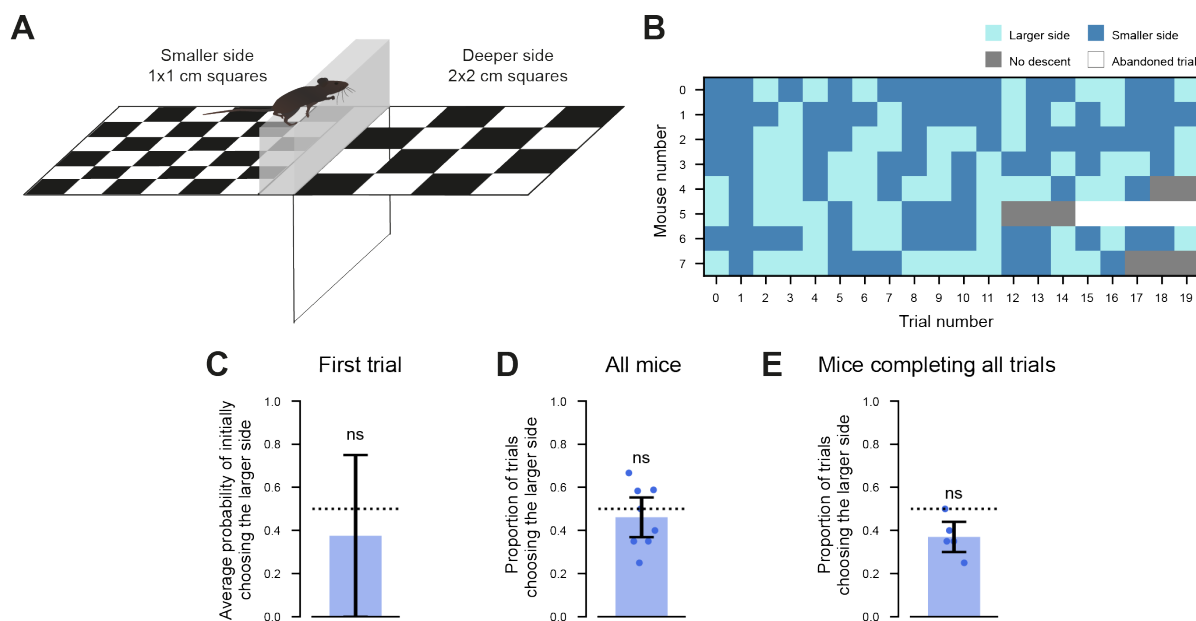


Figure 3.3: Texture density does not bias the decision of mice on the visual cliff. (A) Schematic of visual cliff assay with textures of two densities both on a shallow side. (B) Side of descending for each trial for each mouse on the visual cliff with both textures on the shallow side. (C) Average probability of choosing the larger side on the first trial ($N = 8$ mice). ns: $p > 0.05$. (D) Proportion of trials choosing the larger side (mean $\pm$ 95% confidence interval) for all mice ($N = 8$). ns: $p > 0.05$. (E) Proportion of trials choosing the larger side (mean $\pm$ 95% confidence interval) for mice completing all 20 trials ($N = 5$). ns: $p > 0.05$.

3.6 Depth perception in mice does not depend on prior visual experience

After establishing a baseline performance on the customised visual cliff setup and characterising the effect of the major pictorial cue (i.e. texture density), I then aimed to investigate whether depth perception is innate for mice as demonstrated in many other species. To this end, I reared 23 wild-type mice in complete darkness for 10 weeks to eliminate any prior visual experience before exposing them to the visual cliff. Then I compared their performance on the visual cliff to 17 age-matched control mice reared in normal lighting. All mice were tested on the visual cliff with matched texture density from the perspective of the mice (Figure 3.2A).

When mice were exposed to the visual cliff setup for the very first time, dark-reared mice showed strong preference to the shallow side (86.3%, one sample Wilcoxon test, $p = 1.918 \times 10^{-4}$), and their probability of descending to the shallow side was not significantly different from that of control mice (100%, Figure 3.3C, Wilcoxon rank sum test, $p = 0.1251$). The strong preference to the shallow side at the very first trial after dark-reared mice were exposed to light indicates the presence of depth discrimination ability without any prior visual experience. This is confirmed by the performance of dark-reared mice across the whole session – the average probability of choosing the shallow side for dark-reared mice across the whole session (20 trials) was significantly above chance level (Figure 3.3D-E, all mice: 72.1%, one sample Wilcoxon test, $p = 3.344 \times 10^{-4}$; mice completing all trials: 70%, one sample Wilcoxon test, $p = 9.506 \times 10^{-4}$).

Notably, dark-reared mice still exhibited some differences in performance on the visual cliff compared to control mice. On average, the probability of descending to the shallow side for dark-reared mice across the whole session was lower than that of control mice (Figure 3.4D-E, all mice: Wilcoxon rank sum test, $p = 3.785 \times 10^{-3}$; mice completing all trials: Wilcoxon rank sum test, $p = 3.529 \times 10^{-4}$). Several factors could contribute to this difference. First, dark-reared mice may have higher stress level compared to control mice when they were exposed to light for the first time. During my experiments, the dark-reared mice were more likely to not descend to any side at all compared to the control mice. Among 23 dark-reared mice, 1 never descended and 6 stopped descending after a few trials, whereas only 1 out of 17 control mice stopped descending in the middle of the session (Figure 3.4A-B, F). Approximately one-third of the dark-reared mice took significantly longer to make a decision on the first trial compared to the longest decision time in the control group (276 s, Figure 3.4G). The increased likelihood of no-descend trials and the prolonged decision-making time may suggest heightened stress levels in dark-reared mice when exposed to light. Nevertheless, it is worth noting that even for mice that stopped descending in the middle of a session, most of them still preferred the shallow side (Figure 3.4A,D), which suggests that they were able to discriminate depth despite the high stress level. Second, dark-reared mice may have not fully developed their

motor coordination skills due to the lack of visuomotor feedback during development. The impairment in motor coordination in dark-reared mice has been reported by Lashley and Russell (1934). In my experiments, dark-reared mice exhibited slower movement and reduced exploration of the arena. They also displayed less coordinated behaviours, including shaky posture, difficulty in maintaining balance, and occasional slipping from the central platform. While being blinded to the mouse identity, I re-evaluated the video recordings of each mouse's first trial and classified whether they belonged to the dark-reared or control group. 76% of the dark-reared mice were correctly classified, indicating noticeable differences in movement patterns between dark-reared and control mice (Figure 3.4H).

Overall, these results indicate that depth perception in mice does not depend on any prior visual experience and is thus innate for mice. It also established mice as a reliable and meaningful model for studying depth perception and exploring the underlying genetic, connectomic, and functional mechanisms.

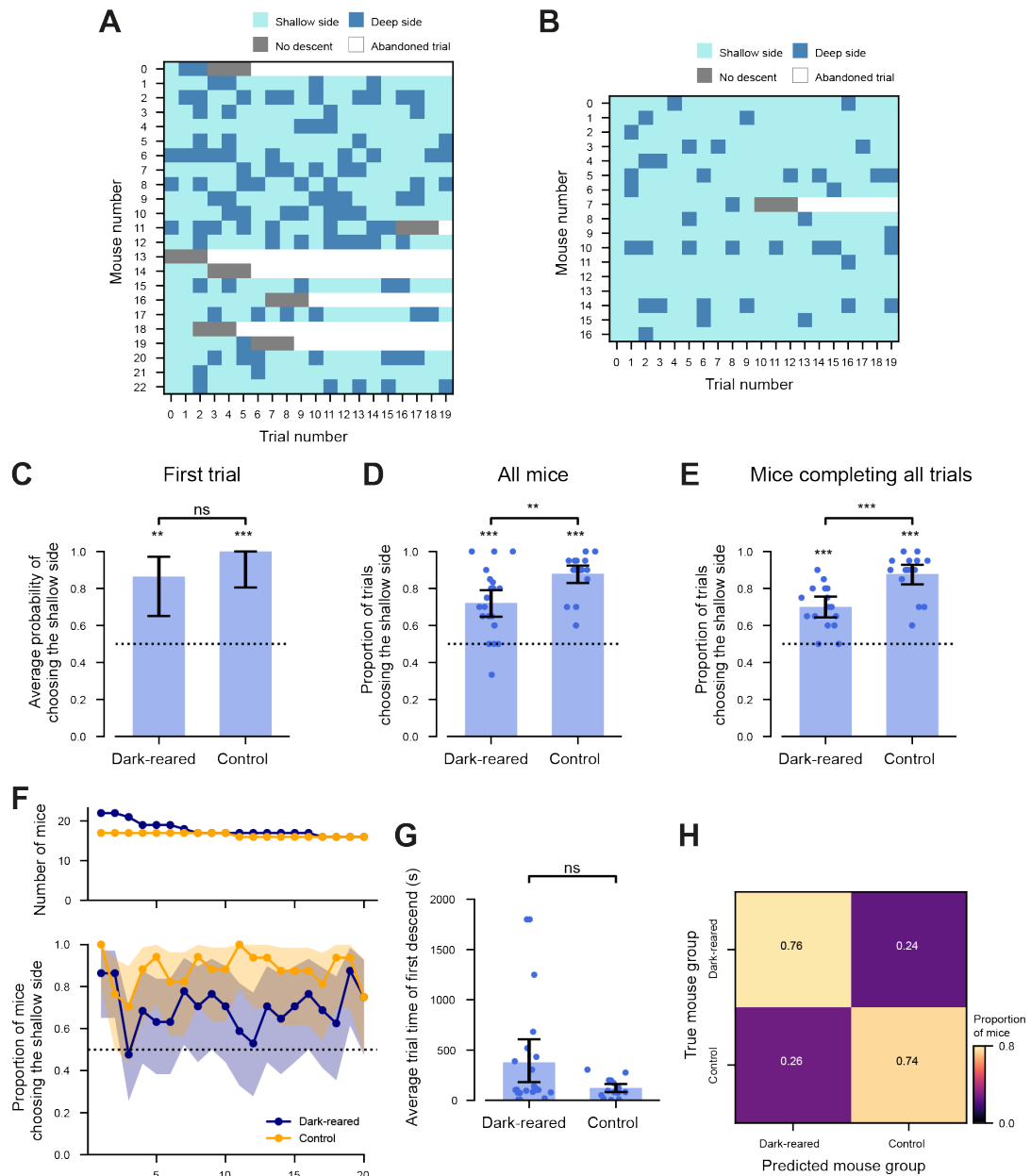


Figure 3.4: Depth perception in mice does not depend on visual experience. (A) Side of descending for each trial for each mouse on the visual cliff for dark-reared mice. (B) Side of descending for each trial for each mouse on the visual cliff for control mice. (C) Average probability of choosing the shallow side ($\pm$ 95% confidence interval) on the first trial of visual cliff test for dark-reared (N = 22) or control mice (N = 17). (D) Proportion of trials choosing the shallow side for all dark-reared (N = 22) or control mice (N = 17). (E) Proportion of trials choosing the shallow side for dark-reared or control mice (N = 16 for each group) who have completed all 20 trials. (F) Bottom - Probability of choosing the shallow side (mean $\pm$ 95% confidence interval) for each trial in dark-reared for control mice. Top - number of mice completing that trial. (G) Average time of the first descend (mean $\pm$ 95% confidence interval) for dark-reared (N = 22) or control mice (N = 17). **: $n < 0.01$, ***: $n < 0.001$, ns: $p > 0.05$. (H) Confusion matrix for the classification of mouse groups (dark-reared or control) based on video recordings of their first trials while being blinded of the identity of the mice.

3.7 Superior colliculus (SC) may not be required for depth perception in mice

To test whether SC activity is required for depth perception in mice, I tested 4 mice on the visual cliff for 20 trials, while bilaterally inhibit their SC activity during half of the trials. These mice expressed Cre-dependent ChRmine (optogenetic mice) in the vesicular GABA transporter-positive (Vgat) cells of their SC via bilateral virus injection. The inhibition of SC was achieved by optogenetically stimulating Vgat cells in SC bilaterally with a ~ 9.8 mW laser at 20 Hz. To control for the non-specific effect of the laser and the surgical procedures, I also tested 4 mice expressing Cre-dependent tdTomato (control mice) on the visual cliff for 20 trials, with laser on for half of the trials and laser off for the other half.

Without laser stimulation, both optogenetic mice and control mice were able to discriminate depth on the visual cliff, choosing the shallow side with a probability of 92.5% and 85% on average respectively (Figure 3.5A-C, hierarchical bootstrap, optogenetic mice: $p = 0.0004$; control mice: $p = 0.0024$, see Methods 2.3.3). Their performance was not significantly different from each other without laser stimulation (Figure 3.5C, hierarchical bootstrap, $p = 0.4568$). This verifies that the *Vgat-cre* mice that I used in this experiment are also able to discriminate depth on the visual cliff, and the surgical procedures including the implantation of optic fibres and virus injections did not affect their depth discrimination ability.

With laser stimulation, the control mice still strongly preferred the shallow side (Figure 3.5B-C, 77.4%, hierarchical bootstrap, $p = 0.0048$), and their performance did not differ significantly compared to their performance without laser stimulation (Figure 3.5C, hierarchical bootstrap, $p = 0.4326$), which suggests that the laser itself did not affect depth discrimination performance on the visual cliff. Under laser stimulation, optogenetic mice also strongly preferred the shallow side (Figure 3.5A,C, 85%, hierarchical bootstrap, $p = 0.0022$), with no significant difference compared to their performance when laser was off (Figure 3.5C, hierarchical bootstrap, $p = 0.32$), which indicates that the inhibition of SC activity does not impair the depth discrimination ability of mice on the visual cliff. Overall, these results suggest that SC is not required for depth perception in mice.

However, the conclusion needs to be interpreted with caution. Firstly, the sample size was quite small ($N = 4$ for each group). Secondly, the efficiency of SC inactivation needs to be further verified. Given the large size of SC, it is challenging to affect the activity of all SC neurons with optogenetics without implanting multiple fibres. In these experiments, the optogenetic approach was designed to achieve inhibition via activating inhibitory neurons, which should result in widespread silencing of the structure. During preliminary experiments (data excluded from this analysis), one mouse exhibited pronounced and frequent head-turning behavioural throughout the 5-s period of stimulation with a continuous-wave laser (as opposed to the pulsed laser used in the presented data).

This observation is consistent with previous reports of SC's role in orienting behaviour (Hoy et al., 2019; Ito & Feldheim, 2018; Shang et al., 2015, 2019; Sparks et al., 1990; Wei et al., 2015). However, the spatial extent and effect of our optogenetic inhibition of SC needs to be further validated with simultaneous in vivo recordings of neuronal activity in SC during laser stimulation.

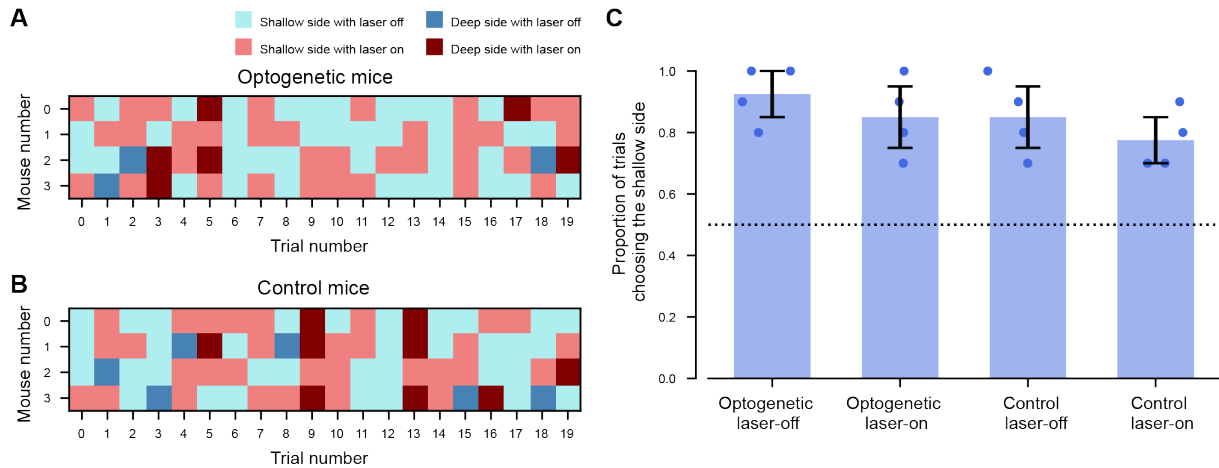


Figure 3.5: **Superior colliculus may not be required for depth perception in mice.** (A) Side of descending for each trial for each mouse on the visual cliff for optogenetic mice. (B) Side of descending for each trial for each mouse on the visual cliff for control mice. (C) Proportion of trials choosing the shallow side (mean $\pm$ 95% confidence interval) for mice completing all 20 trials ($N = 4$ for optogenetic mice, $N = 4$ for control mice).

3.8 Discussion

3.8.1 The nurture and nature of depth perception

Depth perception is an essential visual function that guides animal behaviour in our three-dimensional world. It is important to understand whether this vital visual function is innate for all animals or is developed through visual experience, as an innate behaviour across species implies its functional significance and the presence of evolutionarily conserved neural mechanisms. Previous research have used the visual cliff to study the developmental origin of depth perception. Being presented choices of a shallow side and a deep side with a visual drop ("visual cliff"), most animals with normal vision would choose to descend to the shallow side (Walk & Gibson, 1961). Birds (chicks) as well as many mammals, such as sheep, goats, pigs and rats, were found to show an innate ability to perceive depth on the visual cliff (Gibson & Walk, 1960; Spalding, 1875). Here I adapted the visual cliff setup from Walk and Gibson (1961) and examined whether depth perception is innate for mice. I replicated the preference to the shallow side on the visual cliff observed in Fox (1965) for mice using my custom-built setup. I also confirmed that visual cues, especially non-pictorial depth cues, were required for mice to discriminate depth on the visual cliff. Compared to the visual cliff apparatus used in previous studies, my customised setup provides a simple and reproducible solution that facilitates easy texture changes and randomization of the shallow and deep sides. As a result, I was able to manipulate texture densities presented on both sides to control the pictorial cues displayed, which was not considered in Fox (1965). Finally, by rearing mice in the dark, I show that depth perception of mice does not require any prior visual experience.

It is not surprising that the mouse visual system can support a visual function that does not require prior visual experience, as several parts of the visual system including SC and V1 can already show some mature-like responses at the time of eye-opening for mice (Shen & Colonnese, 2016; Tan et al., 2022; Thompson et al., 2017; Wang et al., 2010). The innate nature of depth perception for mice along with many other species (Gibson & Walk, 1960; Walk & Gibson, 1961) highlights its functional significance in guiding behaviour. It also indicates that the neural circuits supporting depth perception are likely conserved throughout evolution for these species. Therefore, studying the neural mechanisms of depth perception in mice is highly valuable, as it is likely that we are able to extrapolate our findings to other mammals. This innate visual function can also serve as a bridge between behavioural functions, evolutionarily conserved neural circuits and molecularly distinct neuronal populations.

However, environment or visual experience may still play a role in the development or refinement of depth perception. For example, aquatic turtles were able to jump off a raised edge while their terrestrial counterparts refused the action (Yerkes, 1904), demonstrating how distinct environments in which animals have evolved can affect their depth perception or behavioural responses to depth. For kittens, the development of normal depth

perception requires self-initiated active exploration in the environment coupled with visual flow feedback, whereas kittens raised in an environment with only passive visual inputs could not perform depth discrimination (Hein et al., 1970). Some animals exhibit impaired depth discrimination when their visual experience has been deprived through dark-rearing. Kittens reared in the dark for more than 3 months displayed deficiency in visually-guided behaviour such as obstacle avoidance (Crabtree & Riesen, 1979). Infant monkeys reared in the dark for more than a month failed to discriminate depth on the visual cliff (Fantz, 1965). Rats reared in the dark for more than 5 months requires 2 - 4 weeks of visual experience before displaying depth discriminating ability on the visual cliff (Walk et al., 1965). However, dark-rearing may not only abolish visual experience but also impair the normal functioning of neurons in the visual system, leading to impairment in all visually-guided behaviours. Studies have shown that dark-rearing can reduce light-evoked responses in mouse RGCs (Tian & Copenhagen, 2001), lead to poorer binocular responses in neurons of the mouse visual cortex (Tan et al., 2021), alter short-term plasticity of L2/3 pyramidal cells in the rat visual cortex (Liu et al., 2012), and disrupt or delay the maturation of visual cortical functions (Gianfranceschi et al., 2003; Liu et al., 2012; Morales et al., 2002; Müller, 1990). Dark-rearing may also disrupt the development of motor skills, hindering animals' ability to perform behavioural tasks. As an example, our mice raised in the dark for 3 months exhibited impaired motor coordination and high stress level, which may be a potential reason for suboptimal performance on the visual cliff. Therefore, the results of visual function impairments caused by dark-rearing should be interpreted with caution.

It is unclear how much visual experience is required for human infants to develop depth perception as the youngest tested infants were 4-month old (Birch et al., 1982; Fox et al., 1980; Tsuruhara et al., 2014). It is possible that since higher mammals such as primates and carnivores have prolonged period of juvenile growth before adulthood compared to rodents and birds (Bogin, 1997; Paquette & Bigras, 2018), depth perception is not immediately needed after birth while the animals are under the care of their parents. Therefore, while birds and rodents have innate ability of perceiving depth, the maturation of depth perception for primates and carnivores may require some exposure to visual experience during their critical developmental period.

3.8.2 Neural substrates for depth perception

Causal manipulations such as lesions or optogenetic suppression have demonstrated that the visual cortex is indispensable for depth perception in rodents such as rats, mice and Mongolian gerbils (Ellard et al., 1986; Meyer et al., 1966; Parker, Abe, Beatie, et al., 2022). Lesions in the visual cortex of rats impair their performance on the visual cliff (Meyer et al., 1966), and lesions in the visual cortex of gerbils significantly reduces the accuracy of jumping across a gap (Ellard et al., 1986). Optogenetic suppression of the binocular or monocular zone of V1 also impaired the jumping performance of mice under

binocular conditions or monocular conditions, respectively (Parker, Abe, Beatie, et al., 2022).

However, an innate function like depth perception may be supported by subcortical areas that are more evolutionarily conserved. Apart from the image-forming pathway involving dLGN projections to V1, SC controls an important alternative pathway conveying non-image-forming information to V1 (Froudarakis et al., 2019; Juavinett et al., 2020; Seabrook et al., 2017). Receiving the majority of the RGC outputs, SC is key to visuo-motor transformations and visually guided orienting behaviour (Hoy et al., 2019; Ito & Feldheim, 2018; Shang et al., 2015, 2019; Sparks et al., 1990; Wei et al., 2015). Although previous studies have shown that lesions in SC do not impair depth perception performance in tasks such as the visual cliff or jumping across a gap for rats and Mongolian gerbils (Cheney & Crow, 1969; Ellard et al., 1986), it is unknown whether SC is required for depth perception in mice. Here I tested the performance of mice on the visual cliff under optogenetic inhibition of SC. Optogenetic inhibition of SC did not affect the depth discrimination ability of mice on the visual cliff, indicating that similar to other rodents, SC may not be required for depth perception in mice. However, a concrete conclusion cannot be drawn here, as I have not verified the inhibitory effect on the neuronal activity in SC by optogenetic stimulation using simultaneous *in vivo* recordings. It is possible that the silencing of SC in my experiments was not strong enough to create an effect due to insufficient laser power, pulsing frequency or duration.

Nevertheless, if SC is truly dispensable for depth perception in mice, the image-forming pathway involving dLGN, V1 and HVAs are likely to act as vital neural substrates for depth computation. It is unclear whether depth computation from depth cues such as binocular disparity or motion parallax happens at the level of dLGN. Although it was thought that most neurons in dLGN represent monocular information, the activity of some dLGN neurons in cats and monkeys can be modulated by visual stimuli in both eyes (Xue et al., 1987; Zeater et al., 2015), indicating the ability to integrate binocular information. In addition, dLGN neurons are able to integrate visual motion and locomotion-related signals (Eriskien et al., 2014; Roth et al., 2016), which may contribute to the computation of motion parallax. To confirm which areas are required for depth perception, further experiments are needed to establish a causal link between depth perception and the activity of dLGN and different visual cortical areas, by comparing the performance of mice on the visual cliff while inhibiting different areas in the visual thalamus and the visual cortex.

3.8.3 Limitations

My design of the visual cliff paradigm still encompass a few limitations. Pictorial cues such as occlusion, luminance contrast and relative size were limited by measures such as hiding the metal feet, creating arrays of ambient light, and scaling the texture density of the two sides. However, pictorial cues may still be present as mice move around the central platform, as the luminance contrast or the relative size of the textures on the two

sides would vary according to the head position of mice. Although it was shown in section 3.5 that texture density played a limited role in determining the choice of descend for mice on the visual cliff, the setup could be further improved by implementing projections of VR environment onto the visual cliff combined with real-time tracking of the head position of the mice in order to fully control the depth cues presented.

It is also unknown whether mice mainly used binocular disparity or monocular motion parallax cues in depth discrimination on the visual cliff. A recent study suggested that our version of visual cliff test would mainly utilise the lower visual field of mice with far less binocular overlap compared to their upper visual field (Boone et al., 2021). To investigate whether monocular cues are sufficient for mice to perform depth discrimination tasks, future experiments could involve suturing one eye to ensure that only monocular visual cues are available.

3.8.4 Summary

In summary, I established the functional significance of depth perception from non-pictorial visual cues for mice by confirming that depth perception of mice is innate and does not require any prior visual experience. In the next 3 chapters, I will explore in detail how the neuronal activity in the visual cortex, particularly V1, supports the computation of motion parallax signals to infer depth.

Chapter 4

The representation of depth from motion parallax in the primary visual cortex of head-fixed mice

4.1 Authorship declaration

I, Yiran He, performed the majority of the experiments and analyses in this chapter. Antonio Colas Nieto performed $\sim 2/3$ of the two-photon recordings in 2 out of 7 GCaMP6f or GCaMP6s mice, and performed some of the two-photon recordings in the 2 mice injected with GCaMP8m virus. Antonin Blot contributed to the detection of monitor frames and the extraction of eye tracking traces. Xavier Cano-Ferrer and George Constantinou (members of the Making Lab at the Francis Crick Institute) built parts of the customized electronics in the behavioural setup, including screen blanking (see Methods 2.4.8), lick detection, and a printed circuit board for synchronising camera and imaging triggers. This chapter built upon findings that I have previously disseminated in a preprint on bioRxiv (He et al., 2024).

4.2 Background

Motion parallax is a powerful monocular cue for depth perception (Rogers & Graham, 1979). When animals translate in the environment, objects at different distances to the observer appear to move at different velocities on the retinae. If the observer fixates their eyes at optical infinity or cannot converge their eyes to focus on a single point like rodents, all parts in the retinal image move in the opposite direction of self-motion, and the speed of visual motion is inversely proportional to the distance of the target object (Miles, 1998; Yoonessi & Baker, 2011). In this scenario, objects closer to the observer appear to move at a faster speed compared to far objects. If the observer fixates at an intermediate distance, objects closer than the plane for fixation will move in the opposite direction of self-motion, whereas objects farther than the plane of fixation will move in

the same direction as the self-motion direction (Miles, 1998; Nadler et al., 2008; Yoonessi & Baker, 2011). In this case, visual motion speed is inversely proportional to the distance between the object and the plane of fixation. Therefore, the relationship between visual motion and self-motion enables depth estimation by comparing the speed of self-motion and the speed of the resulting visual motion (Figure 4.1A-B, Kral, 2003; Rogers and Graham, 1979).

Many species including humans, non-human primates, rodents and insects are able to make depth judgements based on motion parallax (Cao & Schiller, 2002; Ellard et al., 1984; Kim et al., 2015a; Lehrer et al., 1988; Parker, Abe, Beatie, et al., 2022; Rogers & Graham, 1979; Sobel, 1990a; Srinivasan, 1992). Humans and monkeys can report or detect the shape and corrugation of three-dimensional (3D) surfaces simulated by random-dot stereograms when translating their head or when the stimuli were translated (Cao & Schiller, 2002; Rogers & Graham, 1979; Zhang & Schiller, 2008). Macaque monkeys can also report the depth sign (near or far) from random dot stimuli simulating surfaces at different distances, while being translated horizontally and making compensatory eye movements. Motion parallax is especially important for animals such as rodents who do not have binocular overlaps as extensive as humans and primates (Ellard et al., 1984; Parker, 2007). Several species, including locusts, mongolian gerbils and mice, make translation head movements (head bobs) before jumping across a gap or towards a target, and the number and frequency of head bobbing are positively correlated with the distance required to jump, indicating the use of motion parallax cues for depth estimation (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022; Sobel, 1990a). Optic flow speed simulating a near or far depth can bias the depth judgements of bees during flying (Srinivasan, 1992). These results suggest that motion parallax is an important source of depth information used across many species.

The visual cortex, the primary site for cortical processing of visual information, is a likely neural substrate for depth perception from motion parallax. The visual cortex in mice receives both visual and locomotion-related information from multiple areas including secondary motor cortex (M2), the hippocampal complex, the retrosplenial cortex (RSC), and multiple thalamic nuclei such as LGN, LP and LD (Froudarakis et al., 2019; Juavinett et al., 2020). Neurons in the mouse visual cortex are widely modulated by locomotion, and can integrate visual motion and self-motion information (Ayaz et al., 2013; Dadarlat & Stryker, 2017; Eriskien et al., 2014; Keller et al., 2012; Niell & Stryker, 2010; Saleem et al., 2013). Running significantly amplifies visual responses of excitatory neurons in V1 (Eriskien et al., 2014; Niell & Stryker, 2010). When mice were running in a VR corridor with grating and plaid patterns, V1 neuronal responses can be described by a linear integration between running speed and optic flow speed (Saleem et al., 2013). The ability to integrate visual motion and self-motion information makes the visual cortex the perfect platform to compute motion parallax signals for depth perception.

How neuronal activity the visual cortex supports depth perception from motion par-

allax remains under-explored. Previous studies investigating the neural mechanisms underlying motion parallax in primates typically used passive translations of animals to generate visual motion. Studies in macaques have found that when passively translating the macaques in front of a random dot stimuli, a large proportion of recorded neurons in a higher visual area (HVA) MT display selective responses to specific depths that could be inferred from motion parallax cues (Kim et al., 2015a; Nadler et al., 2008, 2009). However, vision in an active sense in the natural world – animals usually acquire visual information during active exploration of the environment to achieve behavioural goals (Skyberg & Niell, 2024). It has been shown that active locomotion and the resulting closed loop visual feedback is indispensable for cats to develop normal depth perception (Held & Hein, 1963). It is unclear how motion parallax signals are processed by the visual cortex to support depth perception during active navigation of the animals. In contrast, research in rodents often involves naturalistic behavioural tasks, offering less control over the visual stimuli presented. Studies in mongolian gerbils and mice approached depth perception by assessing the animals' ability to jump across a gap (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022). They found that optogenetic suppression of neuronal activity in the monocular zone of the primary visual cortex (V1) significantly disrupted the jumping performance for monocular mice. However, controlling the visual stimuli and depth cues experienced by the animals during freely-moving behaviours such as the jumping task is highly challenging.

Therefore in this chapter, I will present the design of a visual stimulation environment in order to systematically characterise neuronal responses to motion parallax depth cues during active locomotion of animals. I developed a virtual reality (VR) environment with visual cues at different virtual distances to mice while the head-fixed animal was actively exploring the environment. VR environments enable precise control over the parameters of visual stimulation, such as size, location, colour and texture. In addition, the closed loop feedback between the VR environment and the movement of the head-fixed animal offers an opportunity to study motion parallax during active locomotion of animals. It has been demonstrated that mice can perform spatial memory tasks in VR environments (Chen et al., 2018). Neurons critical to navigation in real-world environments, such as place cells, grid cells and head-direction cells, are also active in VR (despite some differences in firing patterns, Chen et al., 2018). Therefore, VR can simulate a sensation of space for mice and serve as a powerful tool to study active vision of mice while allowing precise control over visual stimulation. In this project, I focus my investigation on L2/3 excitatory cells in V1 due to the well-characterised, widespread locomotion-related modulation of V1 neurons (see Section 1.2.4) and the constraints on imageable cortical depth by two-photon microscopy (Yildirim et al., 2019).

4.3 Manipulating virtual depth using motion parallax in head-fixed mice

To investigate whether depth from motion parallax is represented by neurons in V1, I first designed visual stimuli where motion parallax was the only cue for discriminating depth. Ideally, the visual stimuli should provide a 3D experience with easily manipulable parameters. Additionally, the recording setup should be suitable for in vivo two-photon calcium imaging during presentation of visual stimulation while allowing the animals to actively navigate within the visual environment. Two-photon calcium imaging was selected as the recording technique as it covers a larger field of view and records spatial information of neurons compared to in vivo electrophysiological techniques (e.g. tetrode recordings, O’Keefe and Recce, 1993; Wilson and McNaughton, 1993 or neuropixels Steinmetz et al., 2021). Therefore, I chose to establish a recording setup incorporating a 3D virtual reality (VR) environment coupled with a two-photon microscopy. Head-fixed mice could navigate in the VR environment where visual stimuli were presented at different virtual distances, while their neuronal activity in V1 could be imaged by two-photon microscopy (Figure 4.1C).

The 3D VR environment was presented on four monitor screens surrounding the mice, covering ~ 240 degrees of their visual field horizontally and ~ 80 degrees vertically (Figure 4.1C). Mice were head-fixed on a cylindrical wheel and were allowed to freely run on the wheel. The position of the mice in the VR environment was updated in closed loop according to the distance travelled by the mice on the wheel. The VR environment was rendered in the Bonsai software (Lopes et al., 2015). In the virtual environment, mice ran across a textured floor against a gray background, while black spheres at different virtual distances to the mice were presented in their monocular visual field to eliminate binocular cues (Figure 4.1C-G). The locations of the spheres were determined by cylindrical coordinates (Figure 4.1F). The direction of animal’s travel was aligned with the longitudinal axis of the cylinder, and the radius of the cylinder is equivalent to the virtual depth of each trial chosen pseudorandomly from a list of logarithmically spaced values (either 5 virtual depths ranging from 6 cm to 600 cm, or 8 depths ranging from 5 cm to 640 cm). The spheres were spaced along the axis of the animal’s direction of travel, and the azimuth of the spheres were randomly chosen from a uniform distribution between -40 and 40 degrees on the right-hand side of the animal or between 140 and 220 degrees on the left-hand side of the animal (Figure 4.1E). Spheres of the same virtual depth continued to appear until the end of a trial, marked by the time when the animal had travelled 6 m (Figure 4.1F). A 10 s inter-trial interval with only the gray background followed the end of each trial, during which the mice received a probabilistic soy milk reward in 60 - 80% of the trials to motivate their running (Figure 4.1E). The size and density of the spheres in each trial were adjusted so that the spheres maintained the same visual angle at different virtual depths and appeared spaced similarly from the animal’s point

of view (Figure 4.1D,G). The spheres were coloured in solid back without specular highlights (Figure 4.1C,G), eliminating neuronal responses to changes in specular highlights as spheres moved across the animal's visual field. In this VR environment, mice could not distinguish virtual depths of the spheres if they stayed stationary (Figure 4.1G). They could only estimate depth when they started to move, by comparing their own locomotion speeds to the optic flow speeds generated by locomotion (Figure 4.1A-B,G). When mice ran, optic flow speed depends on the animal's locomotion speed and the virtual depth of the spheres – spheres at a near depth trial would create much faster optic flow compared to spheres at a far depth trial.

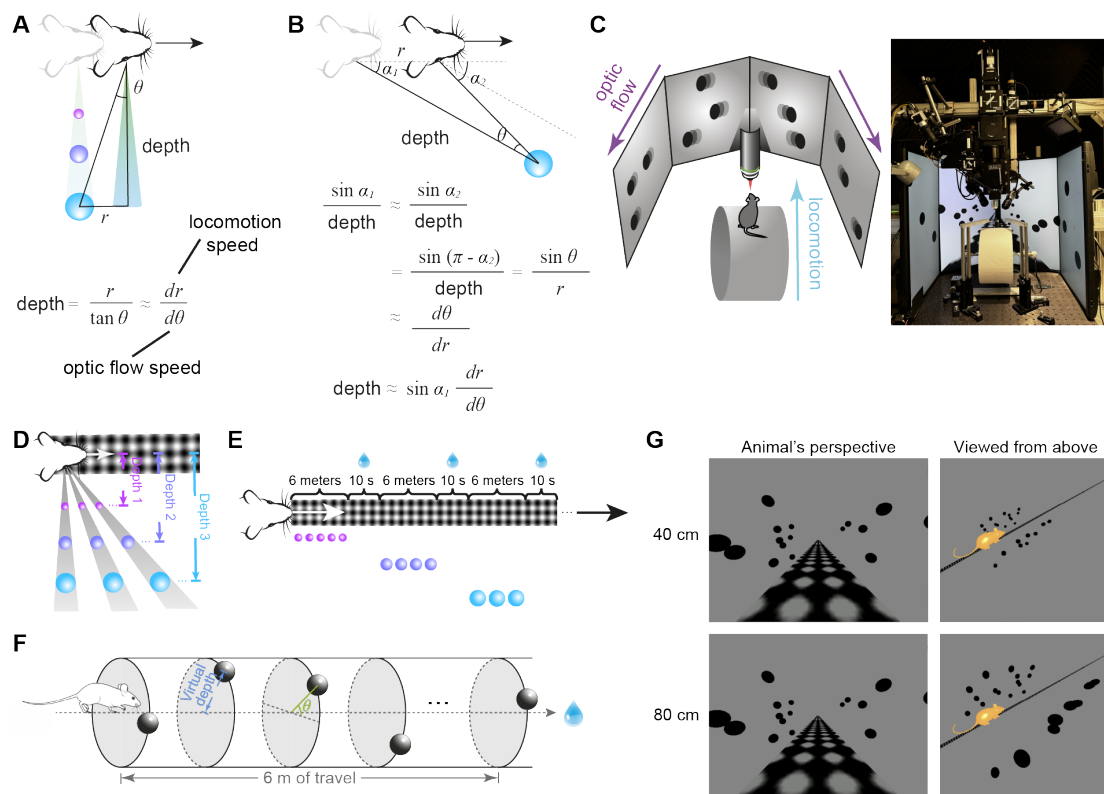


Figure 4.1: Manipulating virtual depth using motion parallax in head-fixed mice. (A-B) Schematic of the relationship between locomotion and visual motion speed due to motion parallax. Depth is the ratio of locomotion speed and optic flow speed when the target object is at 90 degree azimuth (A), and depth is proportional to the ratio of locomotion and optic flow speeds when the object is elsewhere (B). (C) Virtual reality setup and stimulus schematic (left) and a photo of the virtual reality setup (right). (D) Stimulus schematic. Size of spheres at different depths were scaled to cover the same angular extent from the perspective of the mouse. (E) Schematic of trial structure. (F) Schematic illustrating the generation of sphere locations. The locations of the centers of the spheres were selected in cylindrical coordinates. Virtual depth determined the radius of the cylinder, while cylindrical azimuth of the spheres θ (selected from a uniform distribution between -40 and 40 degrees or between 140 and 220 degrees) determined their position relative to the horizon. The mouse drawing from Tyler, E., & Kravitz, L. (2020), Zenodo. doi.org/10.5281/zenodo.3925915. (G) Illustration of stimuli at two virtual depths from the animal's perspective and viewed from above. The mouse (not to scale) is shown to indicate the animals position in the virtual environment.

The four monitors were normally refreshed at 144 Hz. To synchronize the visual stimuli with imaging frames, a small square with varying luminance was presented at the bottom left corner of the leftmost monitor. The square alternated its gray scale at every frame and its luminance changes was recorded by a photodiode, enabling the detection of dropped frames. For 2 mice (25 sessions), the square alternated at every frame between black and white. $89.7 \pm 5.1\%$ of the frame were presented at 144 Hz, resulting in a average frame rate per session above 124 ± 9 Hz. To determine the closed loop latency between rotary encoder inputs and updates of the visual stimulus, for the remaining 5 mice (60 sessions), I updated the brightness of the square following an irregular predefined sequence of 5 values (Figure 4.2A). The sequence was selected to ensure that brightness increases and decreases were alternating at each frame. Frames were detected by the photodiode trace, and the cross-correlation between the filtered predefined sequence and the photodiode recording was computed to identify which frame was presented at each time point and to continuously monitor display lag. The display lag was quantified as the latency between the mouse position measurement by the rotary encoder and the update of the mouse position when rendering the monitor frames in the VR environment, averaged at 26.6 ± 1.4 ms per session (Figure 4.2B-C).

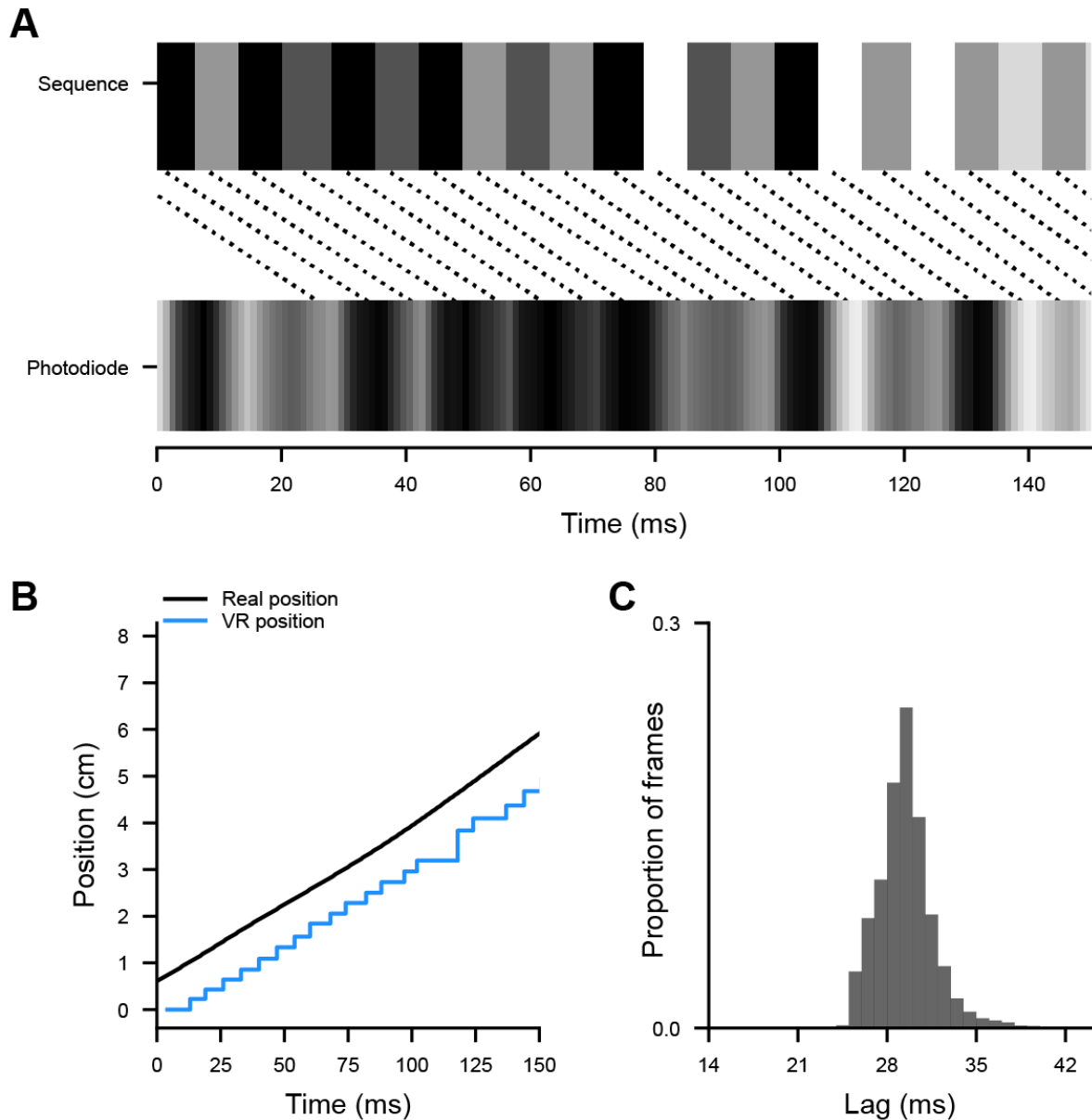


Figure 4.2: **Closed loop synchronisation and latency.** (A) Example sequence of varying luminance presented in the corner of the monitor for synchronization (top) and the recorded photodiode signal (bottom). Dashed line indicates the corresponding frames. (B) Real time mouse position (black), as measured by the rotary encoder, and the current mouse position used to render the VR environment (blue). Each step on the blue trace is a frame and the horizontal shift between the black and blue traces is the closed loop latency. (C) Histogram of VR latencies at each frame.

The positions of the four screens were calibrated using ArUco markers as described in previous literature (Bradski, 2000; Garrido-Jurado et al., 2014, see Methods 2.4.6). In brief, I first estimated the intrinsics of the camera used for calibrating the screen positions by taking pictures of a flat checkerboard pattern at various angles (Figure 4.3A, Kshirsagar et al., 2022). Next, two ArUco markers with known patterns and dimensions were placed in the middle of each monitor screen and the location where the head of mice would be fixed, respectively (Figure 4.3B). After detecting the two ArUco markers in the same picture using the Open CV package (Figure 4.3B, Bradski, 2000), the relative translation and rotation between the two markers could be computed, which indicates the transformation of screens from the perspective of the mice.

Four cameras were used to monitor the eye, facial, and body movements of the mice and were calibrated in a similar way (see Methods 2.4.6). The calibration of the camera monitoring the right eye was shown as an example in Figure 4.3C-D. The intrinsics of the camera were estimated by taking pictures of small checkerboard patterns using the camera from various angles (Figure 4.3C), and the relative rotation from the eye to the camera was estimated using an ArUco marker (Figure 4.3D). As the changes in visual motion speed as a result of gaze direction changes were our primary focus, the translation vector of the eye camera was not relevant here.

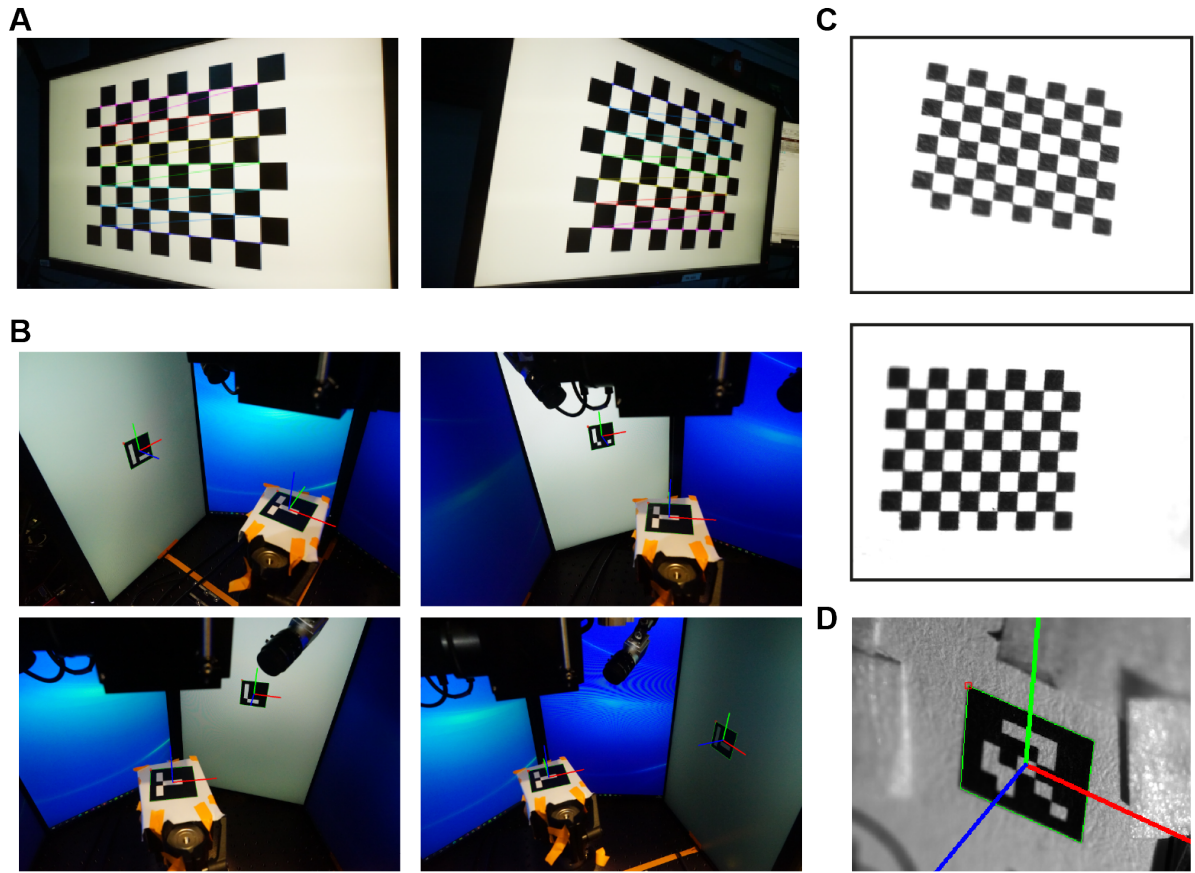


Figure 4.3: **Screen and camera calibration using ArUco markers.** (A) Example photos of a checkerboard pattern at different angles (consists of $3\text{ cm} \times 3\text{ cm}$ black and white squares) taken by the camera used to calibrate screen positions for intrinsics calibration. (B) Example photos of ArUco markers ($5\text{ cm} \times 5\text{ cm}$) displayed both in the centre of each monitor screen and centred at the location where the head of the mice would be fixed at the setup. Photos were taken by the camera in (A). Green lines around the markers show the detection of the marker borders by OpenCV. The red, green, and blue lines pointing out from the centre of the marker indicate the detected x, y, and z axes of the marker respectively. (C) Example photos of a checkerboard pattern at different angles (consists of $0.8\text{ mm} \times 0.8\text{ mm}$ black and white squares) taken by the right eye camera for intrinsics calibration. (D) Example photos of an ArUco marker ($5\text{ cm} \times 5\text{ cm}$) placed parallel to the ground taken by the right eye camera. The borders and axes indicate the marker detection in a similar manner in B.

4.4 L2/3 excitatory neurons in V1 are tuned to depth from motion parallax

To investigate whether V1 neurons selectively respond to specific virtual depths in the VR environment, I recorded neuronal activity in L2/3 of V1 in mice expressing GCaMP6f or GCaMP6s in their pyramidal neurons. Neuronal activity was recorded via two-photon calcium imaging while the mice were navigating the VR environment with spheres at various virtual distances (60,357 neurons from 85 sessions in 7 mice). The VR environment was updated in closed loop with the distance travelled by the animal on the wheel. By visualising the calcium fluorescence ($\Delta F/F_0$) of individual neurons at trials of different virtual depths, I found that many neurons strongly responded to specific virtual depths and were not responsive to others (see examples in Figure 4.4A-B, D-E). To quantify the selectivity towards virtual depth in V1 neurons, I first calculated the average $\Delta F/F_0$ of individual neurons at trials of different virtual depths (Figure 4.4C,F). Then, I fitted the trial average $\Delta F/F_0$ (f) as a function of log-transformed virtual depth (d) with a Gaussian model (see Methods 2.4.11, Figure 4.4C,F):

$$f = ae^{\frac{-(d-d_0)^2}{2\sigma^2}} + b, \quad (4.4.1)$$

where a was the peak response amplitude, σ was the tuning width, and b was the baseline fluorescence. The centre of the fitted Gaussian curve (d_0) represented the log-transformed virtual depth that would triggered the peak response of a neuron, and the preferred depth of that neuron was defined as the exponential of d_0 . To determine whether a neuron is significantly selective for virtual depth, I fitted the Gaussian model above using 5-fold cross validation. On each fold, 80% of the trials were assigned to the training set to estimate model parameters, and the remaining 20% were assigned to the test set to evaluate model predictions. I then calculated the Spearman's correlation between the predicted and observed trial average $\Delta F/F_0$ on test trials across all 5 folds. Depth-selective neurons were defined as having a p-value of Spearman's correlation < 0.05 and a Spearman's correlation coefficient > 0.1 .

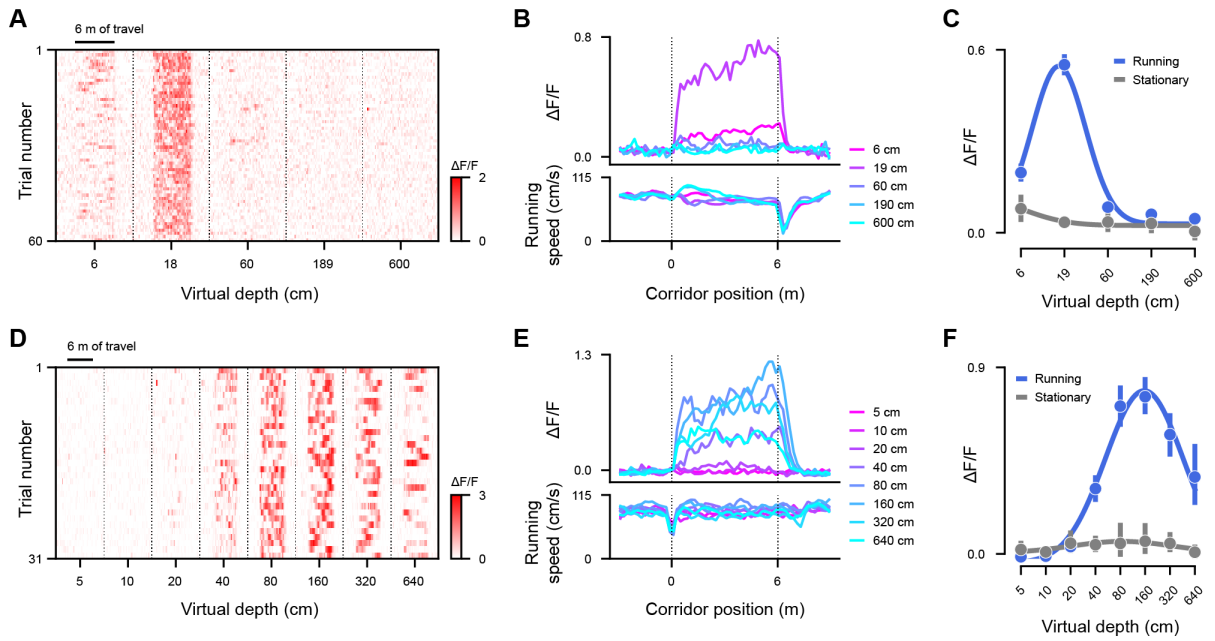


Figure 4.4: **Quantifying depth selectivity for two example neurons.** (A,D) Raster plot of responses of two example neurons in mouse V1 across virtual depths from a session with 5 virtual depths (A) or 8 virtual depths (B) as a function of distance traveled during the trial including 3 m of inter-stimulus interval before and after each trial. Dashed lines separate different trial types. (B, E) Mean responses of the example neuron (top) and mean running speed (bottom) across 5 virtual depths (B) or 8 virtual depths (E). (C, F) Virtual depth tuning of the neuron in A,D during locomotion (>5 cm/s) or stationary (maximum speed during preceding 1 second <5 cm/s) periods. Error bar – 95% confidence interval.

I found that depth selectivity was widespread among V1 L2/3 excitatory neurons, with individual cells tuned to distinct virtual depths (Figure 4.5A-B). Across all recording sessions, a large proportion (51.4%, 31,013 out of 60,357) of L2/3 excitatory neurons displayed significant depth selectivity. The median proportion of neurons in individual recording sessions classified as depth-selective was 33.8% (Figure 4.5C). The preferred virtual depths of depth-selective neurons covered the entire range of depth values probed with the VR stimuli (Figure 4.5D-F).

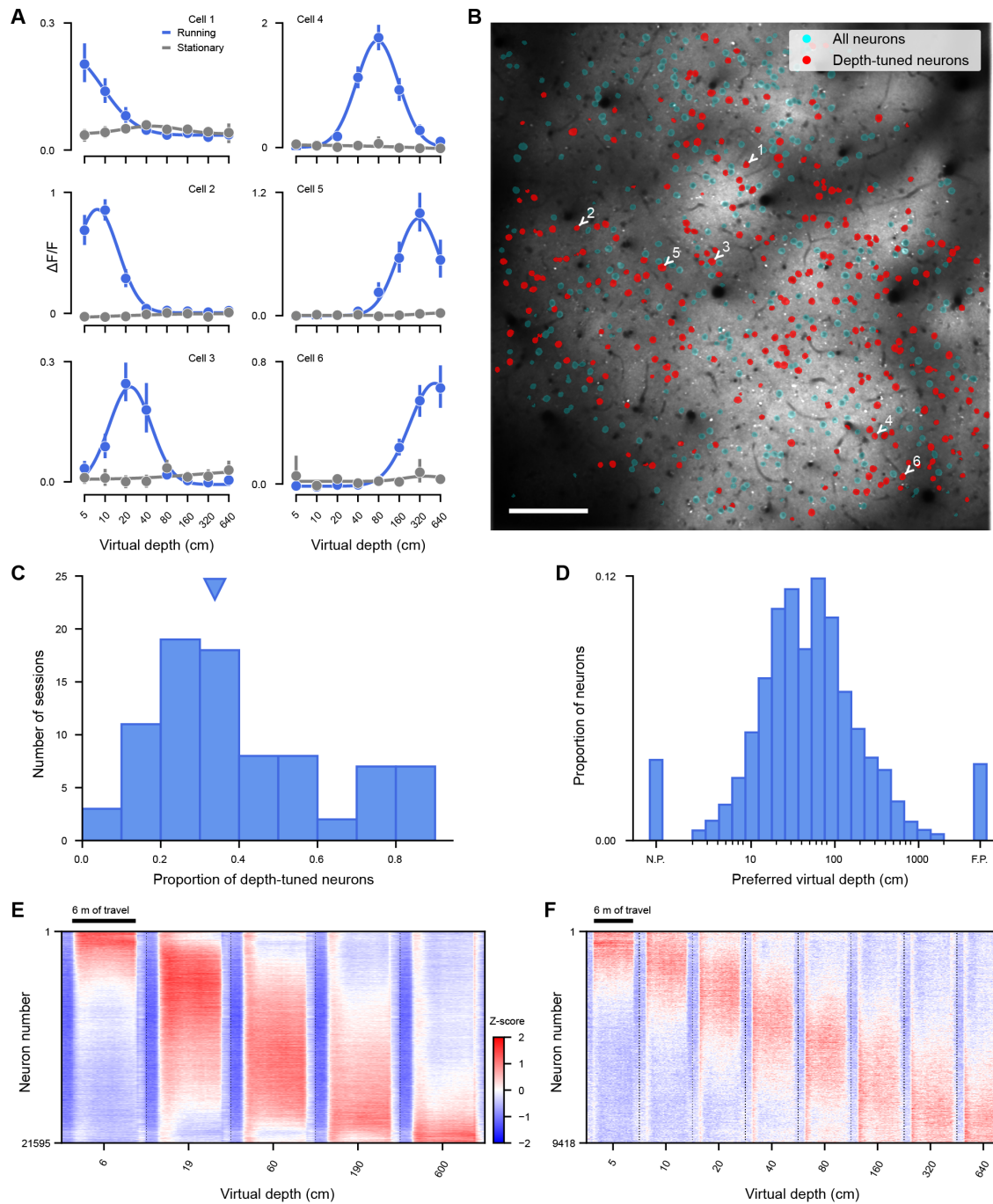


Figure 4.5: Depth selectivity is widespread among V1 L2/3 excitatory neurons. (cont.)

Figure 4.5: **Depth selectivity is widespread among V1 L2/3 excitatory neurons.** **(A)** Virtual depth tuning of 6 additional example neurons during running (blue) and stationary (gray) periods and spatial location of depth-tuned neurons in an example imaging session. Error bar – 95% confidence interval. **(B)** The spatial distribution of depth-selective neurons and all other neurons in the example imaging session, and the locations of example neurons in **(A)**. Scale bar – 100 μm . **(C)** Proportion of depth-tuned neurons across imaging sessions ($N = 85$ sessions). Triangle indicates the median value. **(D)** Distribution of preferred virtual depths of depth-selective neurons ($N = 31,013$ neurons). N.P., near-preferring neurons; F.P., far-preferring neurons. **(E-F)** Raster plot of z-scored responses of all depth-selective neurons from sessions with 5 virtual depths **(E)** or 8 virtual depths **(F)** in mouse V1, in a $[-1 \text{ m}, +1 \text{ m}]$ window around stimulus presentation. Neurons are sorted by virtual depth preferences estimated using hold-out data not used in calculation of the raster plot.

When comparing recordings in GCaMP6f and GCaMP6s mice, a significant difference was observed in the proportion of depth-selective neurons detected, with GCaMP6s sessions showing a significantly higher proportion. To systematically characterise the effect of calcium indicator type on the proportion of depth-selective neurons detected, I compared imaging sessions from GCaMP6f mice ($N = 6$ mice) with those from one GCaMP6s mouse and two additional mice injected with GCaMP8m virus. Sessions using brighter indicators (GCaMP6s and GCaMP8m) showed significantly higher proportions of depth-selective neurons (Figure 4.6B-C, median of GCaMP6s sessions: 81.9%; GCaMP8m: 73.1%) compared to GCaMP6f sessions (Figure 4.6A, median: 31.6%). This likely reflects higher signal-to-noise ratios of GCaMP6s and GCaMP8m indicators (Chen et al., 2013; Zhang et al., 2023), which can be further validated by comparing baseline noise level and neuronal activity across calcium indicators in my experiments. The slightly lower proportion observed in GCaMP8m sessions relative to GCaMP6s sessions may be attributed to the variation in viral expression level using different methods (viral injections vs. genetic encoding) as well as the smaller sample size (6 vs. 14 sessions). Given that the detection of depth-selective neurons can be limited by the signal-to-noise ratio of calcium indicators, the percentage of depth-selective neurons reported here likely represents a lower bound of the true prevalence of depth-tuned neurons in the brain.

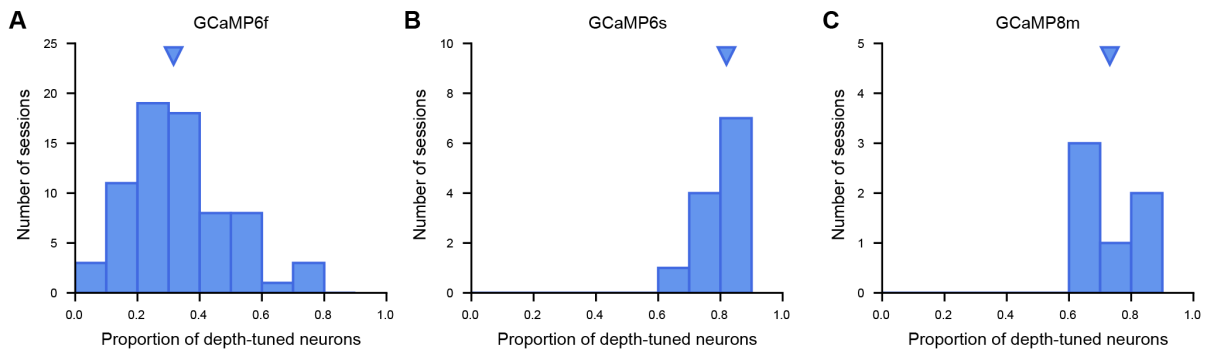


Figure 4.6: Depth-selective neurons measured by different calcium indicators. (A-C) Proportion of depth-tuned neurons across imaging sessions recorded in GCaMP6f mice (A, $N = 71$ sessions from 6 mice), the GCaMP6s mouse (B, $N = 14$ sessions from 1 mouse), and wildtype mice injected with GCaMP8m virus (C, $N = 6$ sessions from 2 mice). Triangle indicates the median value.

Moreover, neurons did not show depth-selective responses when mice were stationary which was defined by moments when the maximum running speed in the preceding 1 second was lower than 5 cm/s (Figure 4.5A). To examine whether running improves the encoding of depth by V1 populations, I trained a linear support vector machine (SVM) decoder to decode virtual depth from population activity of V1 neurons in each recording session, and then compared the classification errors at frames when mice were stationary or running at different ranges of running speeds. The confusion matrices from the example sessions (Figure 4.7A-B) and the average confusion matrices from all sessions (Figure 4.7C-D) demonstrated that while virtual depth could be accurately decoded when mice were running (e.g. at speeds between 40 - 60 cm/s), it could not be decoded during stationary periods. The decoding error was around chance level during stationary periods, and was significantly reduced when mice started to run above 5 cm/s (Figure 4.7E). The results show that the representation of depth in mouse V1 is driven by motion parallax as a result of animals' locomotion.

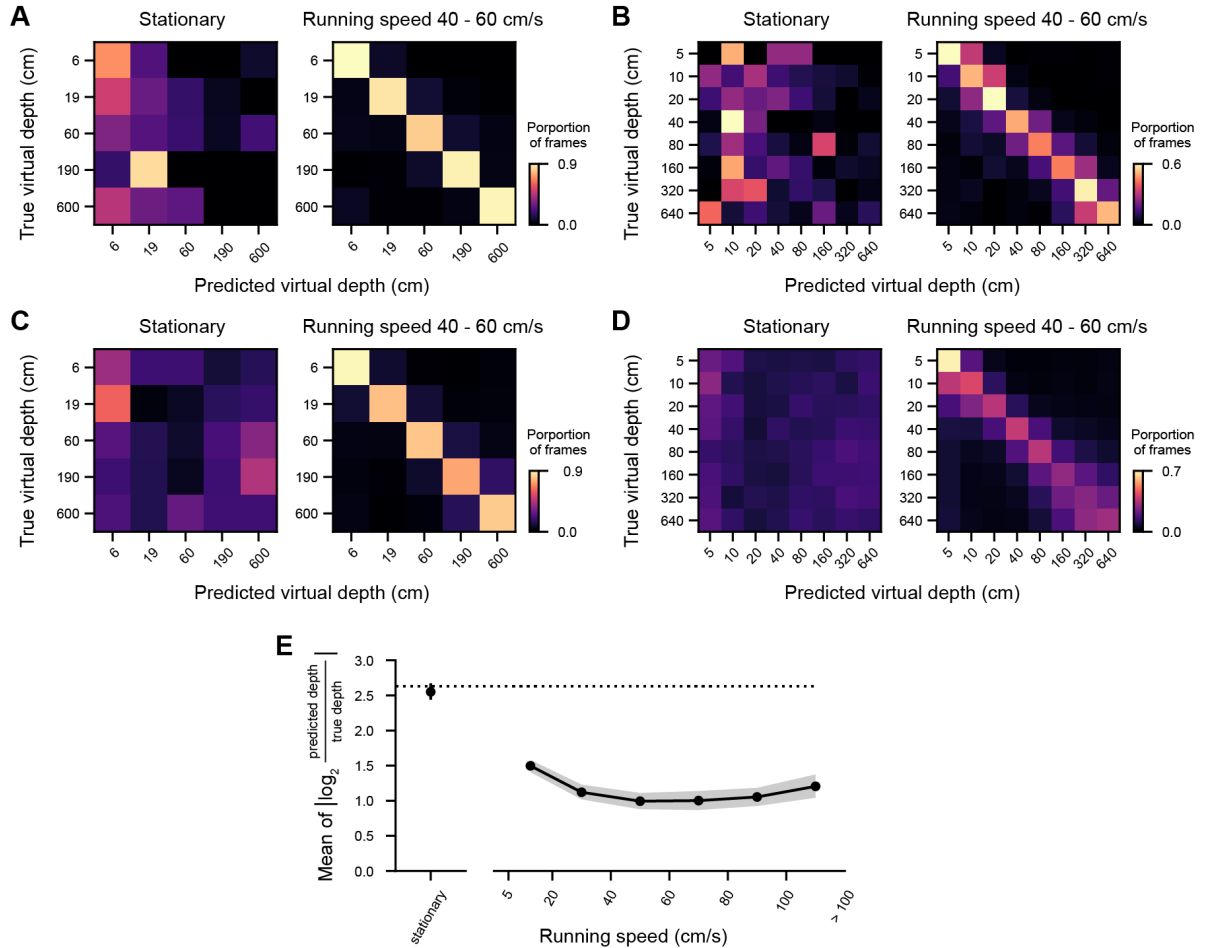


Figure 4.7: Depth selectivity in V1 is driven by motion parallax. (A-B) Confusion matrices of decoded virtual depth from an example session with either 5 (A) or 8 virtual depths (B), when animals were stationary or running at a speed of 40-60 cm/s. (C-D) Average confusion matrices of decoded virtual depth when animals were stationary or running at a speed of 40-60 cm/s, from all sessions with either 5 virtual depths (C, N = 25 sessions) or 8 virtual depths (D, N = 60 sessions). (E) Mean error of virtual depth decoded from population activity of simultaneously recorded neurons when mice were stationary and as a function of running speed (N = 85 sessions). Dashed line – chance level; shading – 95% confidence interval.

As neuronal activity in mouse V1 is highly modulated by behavioural states (Miura & Scanziani, 2022; Saleem et al., 2013; Stringer et al., 2019), it is necessary to confirm whether systematic differences in behaviour could confound our interpretation of the recorded neuronal responses. First, I compared the average running speed across trials of different virtual depths. Across corridor positions in each trial, mice showed a similar movement pattern – they stopped temporarily when the spheres appeared, then ran at a sustained speed throughout the trial, and finally stopped for the reward after trial ended (Figure 4.8A-B). Across trials of different virtual depths, mice ran at similar speeds, despite slight reduction in the farthest depths in the 5-depth condition (Figure 4.8C) or in the nearest depths in the 8-depth condition (Figure 4.8D). It is unlikely that the depth-selective responses I observed were elicited solely by changes in running speed, as individual neurons displayed a vast range of depth preferences. I will further examine the relationship between running speed and depth-selective responses in Section 5.3 in next chapter. Next, I examined whether depth-selective responses might be a result of stereotyped eye movements triggered by stimuli of different depths. Stereotyped eye movements, especially rapid saccades in response to attempted head rotation (Meyer et al., 2020), would change the retinal image and visual motion speeds of the stimuli drastically, confounding V1 responses to virtual depths. To this end, I compared the average eye velocity and saccade rate across trials of different virtual depths by tracking the pupil of the right eye of the mice in a subset of recordings (Figure 4.8I-L). Across trials of different virtual depths, eye velocity and saccade rate were similar (Figure 4.8K-L), indicating that no stereotyped eye movements were triggered. The saccade rate was also very low on average (~ 0.05 Hz, Figure 4.8L), with 1 saccade happening in every two trials based on the average running speed (Figure 4.8C-D). Therefore, the depth-selective responses I observed cannot be attributed to behavioural changes in running speeds or eye movements.

4.5 Depth selectivity is invariant to stimulus size

To maintain a consistent visual angle for sphere stimuli across various virtual depths, the size of spheres in the virtual environment was scaled proportionally to the virtual depth for each trial (Figure 4.1D). Consequently, the observed depth-selective responses could be attributed to selectivity for size of the spheres. To examine this possibility, we recorded the neuronal activity in response to both variation in virtual depth and sizes of the spheres in 3 recordings. At each virtual depth trial, spheres subtended either 5, 10 or 20 degrees of the visual field in diameter when the mice passed closest to each sphere in the VR. As V1 neurons were tuned to stimulus size (Adesnik et al., 2012; Dipoppa et al., 2018), the magnitude of neuronal responses were modulated by the visual angle of spheres as expected (Figure 4.9A-B). However, the preferred virtual depth was maintained in response to spheres of different visual angles (Figure 4.9C). This indicates that depth

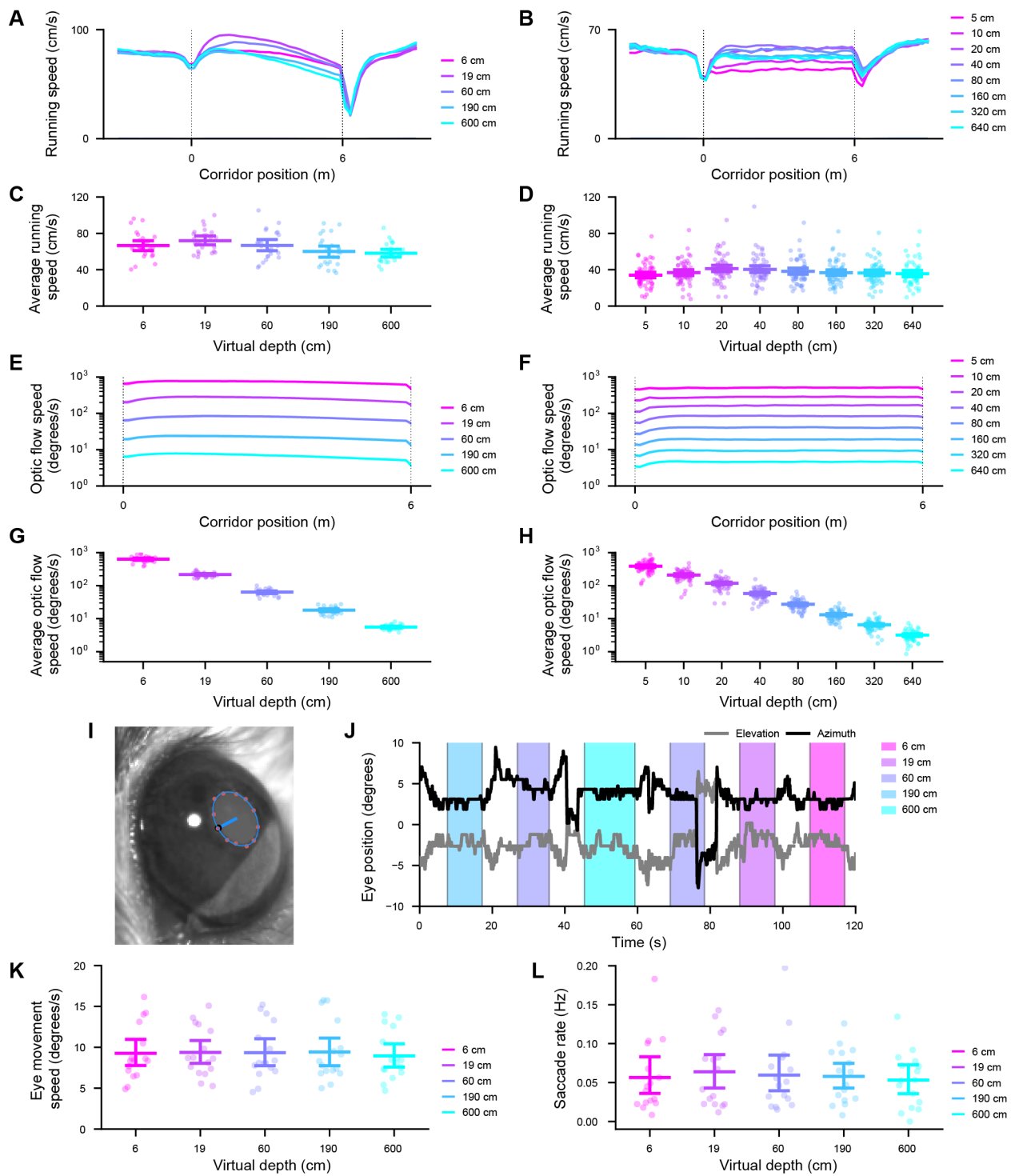


Figure 4.8: Running, optic flow speeds and eye movements across different virtual depths. (*cont.*)

selectivity is unaffected by stimulus size.

Figure 4.8: **Running, optic flow speeds and eye movements across different virtual depths.** **(A-B)** Mean of session average running speed as a function of distance traveled during the trials at 5 (**A**, $N = 25$ sessions) or 8 (**B**, $N = 60$ sessions) virtual depths, including 3 m of inter-stimulus interval before and after the trials. Shading – 95% confidence interval. **(C-D)** Mean running speed of each session at 5 (**C**) or 8 (**D**) virtual depths. Error bar – 95% confidence interval. **(E-F)** Mean of session average optic flow speed as a function of distance traveled during the trials at 5 (**E**) or 8 (**F**) virtual depths. Shading – 95% confidence interval. **(G-H)** Mean optic flow speed of each session at 5 (**G**) or 8 (**H**) virtual depths. Error bar – 95% confidence interval. **(I)** Example frame of the right eye recorded during imaging. The pupil border was tracked using DeepLabCut (orange dot), fitted with an ellipse (blue circle) and gaze (blue line) from the eye center (black dot) was estimated as described in Wallace et al., 2013. **(J)** Example gaze direction in azimuth and elevation relative to median direction during a recording. **(K-L)** Average eye velocity (**K**) and saccade rate (**L**) by session across trials of different virtual depths ($N = 16$ sessions, 2 mice).

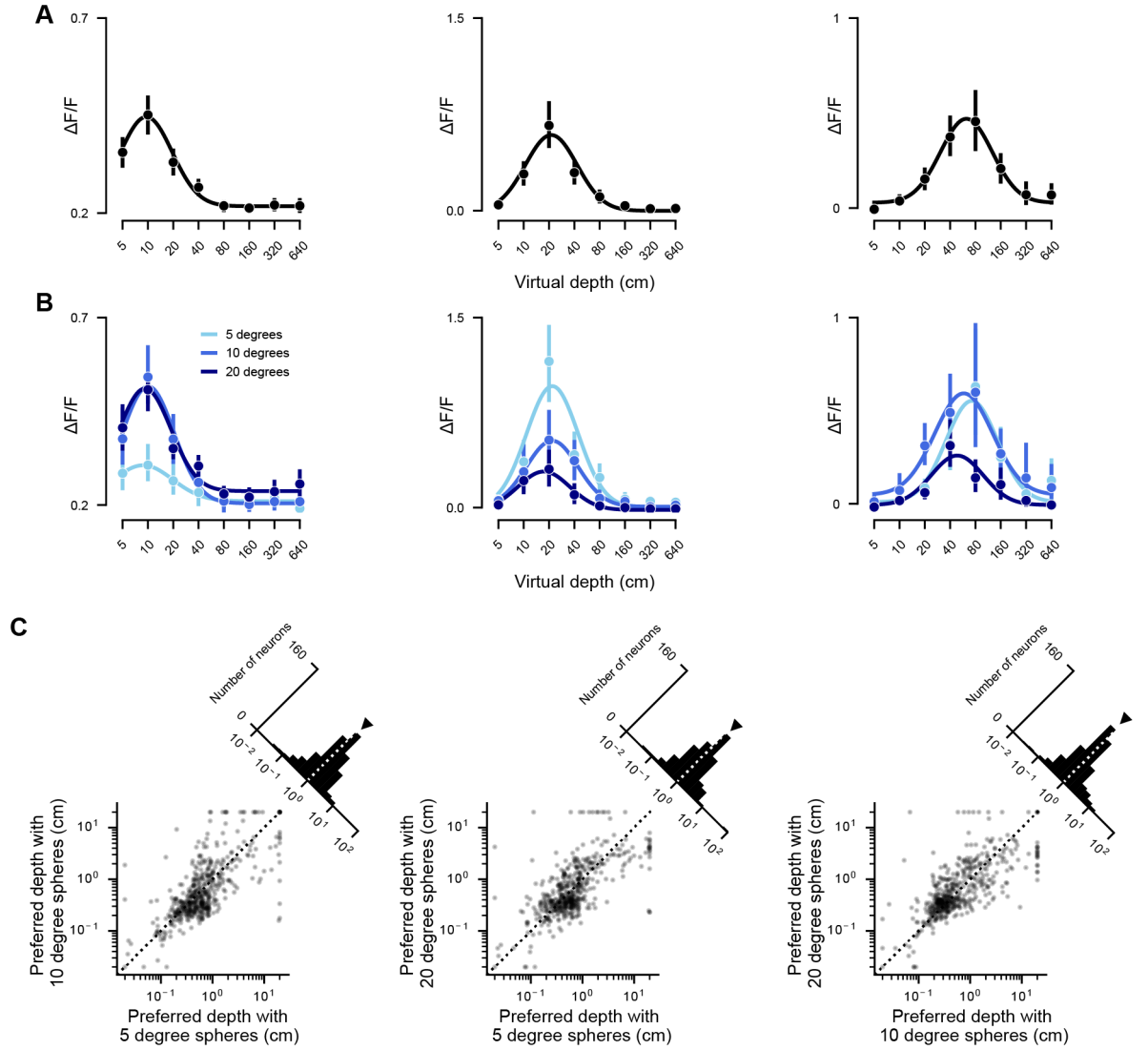


Figure 4.9: **Virtual depth selectivity is invariant to stimulus size.** (A) Virtual depth tuning of 3 example neurons across all stimulus sizes. (B) Virtual depth tuning of the neurons in (A) in response to different sizes of sphere stimuli. (C) Preferred virtual depths of all depth-selective neurons (500 neurons from 3 sessions in 2 mice) across pairs of stimuli. Triangles indicate medians. Insets – histograms of the ratio of preferred virtual depths across pairs of stimuli sizes. Preferred virtual depth mapped with 5 degree vs. 10 degree spheres, median ratio = 1.13, $p_{ratio} = 0.232$, $r = 0.678$, $p_{correlation} < 0.0001$; 5 vs. 20 degree spheres, median ratio = 1.13, $p_{ratio} = 0.401$, $r = 0.635$, $p_{correlation} < 0.0001$; 10 vs. 20 degree spheres, median ratio = 0.93, $p_{ratio} = 0.229$, $r = 0.721$, $p_{correlation} < 0.0001$.

4.6 Discussion

4.6.1 Studying the neural representation of depth from motion parallax using virtual reality

Motion parallax is a key monocular cue for depth estimation (Rogers & Graham, 1979) and is particularly important for many animals without extensive binocular overlap, such as rodents (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022). Animals can infer depth by using vision as an active sense – when moving through the environment, objects at different distances will appear to move at different velocities. However, the neural mechanisms underlying motion parallax during active navigation have not been systematically explored. The visual cortex, especially V1, is a likely neural substrate for depth perception from motion parallax. Optogenetic suppression of the neuronal activity of the monocular V1 significantly disrupts the ability of mice to jump across a gap, suggesting that V1 is critical to depth estimation (Parker, Abe, Beatie, et al., 2022). Additionally, V1 neurons can integrate visual motion and self-motion information (Ayaz et al., 2013; Dadarlat & Stryker, 2017; Erisken et al., 2014; Keller et al., 2012; Niell & Stryker, 2010; Saleem et al., 2013), the two key components required for computing motion parallax. However, existing studies have primarily examined motion parallax either in the context of passive translations (Kim et al., 2015a; Nadler et al., 2008, 2009) or in freely-moving tasks without precise control of visual stimulation (Ellard et al., 1984; Parker, Abe, Beatie, et al., 2022). To address this gap, I aimed to design experiments to systematically characterise how V1 supports depth perception from motion parallax during active locomotion, precise control over the visual stimuli presented to the animal.

In this chapter, I introduced the design of a 3D VR environment to provide motion parallax depth cues for head-fixed mice. The VR setup presented spheres at varying virtual distances to the mice. During active locomotion of head-fixed mice on a wheel, the positions of spheres in the virtual environment were updated according to the running speed of the mice. Consequently, optic flow was generated by active running of the mice. Objects at different virtual depths created optic flow at distinct speeds, equivalent to the ratio between running speed and virtual depth. VR offers a powerful tool to study depth perception from motion parallax. It has been shown that mice can perform spatial memory tasks in VR environments and the neurons supporting spatial perception remain active in VR environments (Chen et al., 2018), indicating a sensation of space for mice in VR. Moreover, VR allows easy and precise manipulation of the visual cues presented, which enabled the isolation of motion parallax as the only available depth cue. For example, objects were presented exclusively in the monocular zone to eliminate binocular depth cues. Spheres were chosen as the visual stimuli to minimise the influence of shape changes with shifts in the observer's perspective. Size of spheres were scaled to maintain a constant visual angle across all virtual depths, preventing depth-selective responses from being confounded by changes in stimulus size. By restricting pictorial and binocular cues,

this design ensured that motion parallax remained the primary cue for distinguishing trials with different virtual depths. Coupling VR environments with active locomotion in head-fixed mice enables us to study the neural mechanisms underlying depth perception from motion parallax during active navigation, while allowing for precise control over visual stimulation.

4.6.2 The representation of depth in V1

Using this 3D VR environment, I recorded L2/3 excitatory neurons in V1 when mice were actively running in the environment. I found that a large proportion (51.4%) of the recorded L2/3 excitatory neurons exhibited significant selectivity to virtual depths of stimuli in VR. Neurons tuned to the full range of the virtual depth values I probed were found in V1, allowing V1 to represent a wide range of distances that are ethologically relevant for mice in the natural environment. I also confirmed that depth selectivity of V1 neurons was driven by motion parallax as a result of active locomotion of the mice on the wheel. Depth-selective responses were absent when animals were stationary, and virtual depth could not be decoded from the population activity during stationary periods. Compared to previous research that investigated motion parallax through passive translations (Kim et al., 2015a; Nadler et al., 2008, 2009), my findings provide new insights into the neural mechanisms underlying depth perception from motion parallax in a scenario more reflective of natural conditions and demonstrate that the visual cortex can represent depth during active navigation of the animal.

The signal-to-noise ratio of the calcium indicator used can hinder the detection of low neuronal responses. When imaging using a calcium indicator with higher signal-to-noise ratios such as GCaMP6s or GCaMP8m (Chen et al., 2013; Zhang et al., 2023), the proportion of depth-selective neurons identified was significantly higher compared to the imaging sessions using GCaMP6f. Although the proportion of depth-selective neurons measured in mice expressing GCaMP8m was not higher than that measured in mice expressing GCaMP6s, this may be attributed to variations in virus expression levels resulting from the virus injection process in mice expressing GCaMP8m compared to transgenic GCaMP6s mice. Therefore, the percentage of neurons with significant depth selectivity is likely to represent a lower bound of the actual proportion of neurons in the brain that are tuned to depth.

Neuronal responses in V1 are modulated by a variety of behaviours such as running, eye movements, pupil dilation and facial movements (Ayaz et al., 2013; Dadarlat & Stryker, 2017; Eriskin et al., 2014; Keller et al., 2012; Niell & Stryker, 2010; Parker, Abe, Leonard, et al., 2022; Saleem et al., 2013; Stringer et al., 2019). In my experiments, two key factors that could potentially affect responses to different virtual depths are running speed and eye movement. If mice exhibited stereotyped running speed changes or eye movements in response to specific virtual depths, the observed depth-selective responses might simply reflect these stereotyped behaviours. To test this, I first compared the average running

profiles of mice across trials of different virtual depths, and confirmed that the running speeds were similar across virtual depths. This indicates that depth-selectivity cannot be solely attributed to differences in running speed. Next, I analysed the eye velocity and saccade rate during trials of different virtual depths. Eye movements, especially rapid saccades, can cause abrupt changes in the visual motion speed, potentially confounding neuronal responses to motion parallax cues. However, no significant differences were found in eye velocity or saccade rate across virtual depths, suggesting that depth-selective responses are unlikely to result from stereotyped eye movement patterns. The eye movement patterns observed in my experiments also aligns with previous research (Meyer et al., 2020), showing that head-fixed mice typically do not exhibit drastic and dynamic eye movements. One potential confounding factor requiring further analysis is facial movement, as certain depths (e.g. very near depth) may trigger facial movements such as whisking, which was found to modulate the spontaneous activity in V1 (Stringer et al., 2019). further analysis is needed using tools such as Facemap (Syeda et al., 2024) to track and quantify the facial movement of the mice. To address this, further analysis could employ tools such as Facemap (Syeda et al., 2024) to track and quantify the facial movements of mice.

4.6.3 Summary

V1 neuronal responses have traditionally been studied using two-dimensional visual stimuli such as gratings or natural images (Andermann et al., 2011; Bashivan et al., 2019; Bonin et al., 2011; Dräger, 1975; Fu et al., 2024; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008; Walker et al., 2019). This has contributed to a widely accepted view that V1 neurons act as a set of Gabor filters over two-dimensional retinal images, exhibiting selectivity for spatial and temporal frequency, orientation and direction (Adelson & Bergen, 1985; Baker & Issa, 2005; Basole et al., 2003; Bonin et al., 2011; Mante & Carandini, 2005; Niell & Stryker, 2008). Here I established a VR setup which enabled systematic characterisation of V1 neuronal responses to motion parallax depth cues when mice were actively navigating in a 3D virtual environment. I found that V1 neurons are tuned to the third-dimension of the visual environment – depth, during active locomotion. How this depth selectivity arises from locomotion and visual motion signals and how the representation of depth is distributed across V1 will be illustrated in the next two chapters.

Chapter 5

The emergence of depth-selective activity from the integration of locomotion and visual motion signals

5.1 Authorship declaration

I, Yiran He, performed the majority of the experiments and analyses in this chapter. Antonio Colas Nieto performed $\sim 2/3$ of the two-photon recordings in 2 out of 7 GCaMP6f or GCaMP6s mice. Xavier Cano-Ferrer and George Constantinou (members of the Making Lab at the Francis Crick Institute) built parts of the customised electronics in the behavioural setup, including screen blanking (see Methods 2.4.8), lick detection, and the printed circuit board for synchronising camera and imaging triggers. This chapter built upon findings that I have previously disseminated in a preprint on bioRxiv (He et al., 2024).

5.2 Background

To compute depth from motion parallax, the brain needs to get access to two critical signals – the speed of self-motion and the speed of visual motion as a result of the self-motion. In mice, eye movements are mainly made to compensate for head rotations (Holmgren et al., 2021; Meyer et al., 2020; Michael et al., 2020). Therefore, body movements such as walking or running are the main source of locomotion to induce motion parallax for mice. Multiple studies have shown that excitatory neurons in mouse V1 can respond to both signals. V1 neurons are tuned to speed of visual motion (Andermann et al., 2011; Marshel et al., 2011). The spontaneous and visual-driven activity in V1 is also highly modulated by the running speed of mice (Keller et al., 2012; Niell & Stryker, 2010; Saleem et al., 2013), indicating that V1 neurons are able to integrate visual motion and self-motion signals (Keller et al., 2012; Saleem et al., 2013).

The functional significance of visuomotor integration in V1 has been a topic of active

debate. Some have proposed that locomotion modulates the gain of sensory responses in V1, thereby enhancing the encoding of visual information within V1 populations (Bennett et al., 2013; Dadarlat & Stryker, 2017; Harris & Thiele, 2011; Niell & Stryker, 2010). Others have suggested that visuomotor integration contributes to predictive coding, where visual responses reflect the prediction error between the predicted and actual sensory feedback to help form an internal model of the world (Jordan & Keller, 2020; Keller & Mrsic-Flogel, 2018; Keller et al., 2012; Muzzu & Saleem, 2021; Zmarz & Keller, 2016). No consensus on the function of V1 visuomotor integration in visual processing has yet been reached. Given that depth can be computed from the ratio of self-motion speed and the resulting visual motion speed, I propose that visuomotor integration in V1 neurons enables the estimation of depth from motion parallax cues.

In this chapter, I aim to test the hypothesis that depth selectivity in V1 demonstrated in the previous chapter arises from the integration of running speed and optic flow speed signals. I will also characterise how the integration occurs in depth-selective neurons by comparing different computational models of running and optic flow speed integration. Finally, I will explore on a population level, whether closed loop coupling contributes to a more accurate representation of depth among V1 populations.

5.3 Depth selectivity arises from conjunctive coding of running speed and optic flow speed

Optic flow speed varied across trials of virtual depths. Optic flow speed generated by locomotion is determined by the ratio of locomotion speed to virtual depth. With similar running speed, nearby depths elicit faster optic flow speed compared to far depths (Figure 4.8E-H). Therefore, depth-selective responses could simply be a result of neuronal selectivity for optic flow speeds. However, depth from motion parallax is ambiguous without extra-retinal inputs – slow visual motion can indicate either a far object or a slow-moving observer. Therefore, both running speed and optic flow speed are important for depth estimation. I hypothesize that depth selectivity of V1 neurons measured in the previous chapter is a product of integrating both optic flow and locomotion-related signals.

To explore the relationship between depth selectivity, running speed and optic flow speed, I first visualised the activity of depth-selective neurons as a function of running speed and optic flow speed (Figure 5.1B,F). Optic flow speed of each imaging frame was calculated as the ratio of running speed and virtual depth, corresponding to the visual motion speed when spheres were at 90 degree azimuth. Two key observations can be drawn from this visualisation. First of all, if depth selectivity can be solely attributed to responses to different optic flow speeds, optic flow tuning should remain the same across virtual depths. On the contrary, for depth-selective neurons, the same optic flow speed resulted in substantially distinct neuronal responses at different virtual depths

(Figure 5.1B,F). Given that different running speeds were needed to generate the same optic flow speed across virtual depths, this suggested that both running and optic flow speed contribute to depth-selective activity. Neuronal responses as a function of both running and optic flow speeds further supported this conclusion (Figure 5.1C,G). If a depth-selective neuron were exclusively tuned to optic flow speed (or running speed), the matrix would show that the neuron is only responsive on a specific horizontal (or vertical) line while maintaining invariant responses to the other axis. Instead, the depth-selective responses I observed are neither a product of optic flow speed tuning or running speed tuning alone, but an integration between these two signals. The second interesting point is that, since virtual depth is defined as the ratio of optic flow speed to running speed, different virtual depths should correspond to the parallel 45-degree lines with varying intercepts on the y-axis in the plot of log-optic flow speed versus log-running speed (in Figure 5.1C,F). Depth-selective neurons seemed to prefer specific conjunctions of running speed and optic flow speed, and the preferred conjunctions lie on the line corresponding to their preferred virtual depths (Figure 5.1C,F). This will be illustrated further in the next section (Figure 5.5A-B).

To further quantify how running and optic flow speed were integrated, I fitted the responses of depth-selective neurons as a function of running speed and optic flow speed using five models: a model where neurons were tuned to either optic flow speed or running speed (optic flow model or running speed model); an additive model where neuronal responses were a linear summation of running and optic flow responses; a conjunctive model where neuronal responses were driven by a specific conjunction of running and optic flow speed tuning modelled by a two-dimensional Gaussian function; and an idealised depth tuning model (RS/OF model) where neurons were tuned to the ratio of running and optic flow speed (Figure 5.1D,H, 5.2A-B).

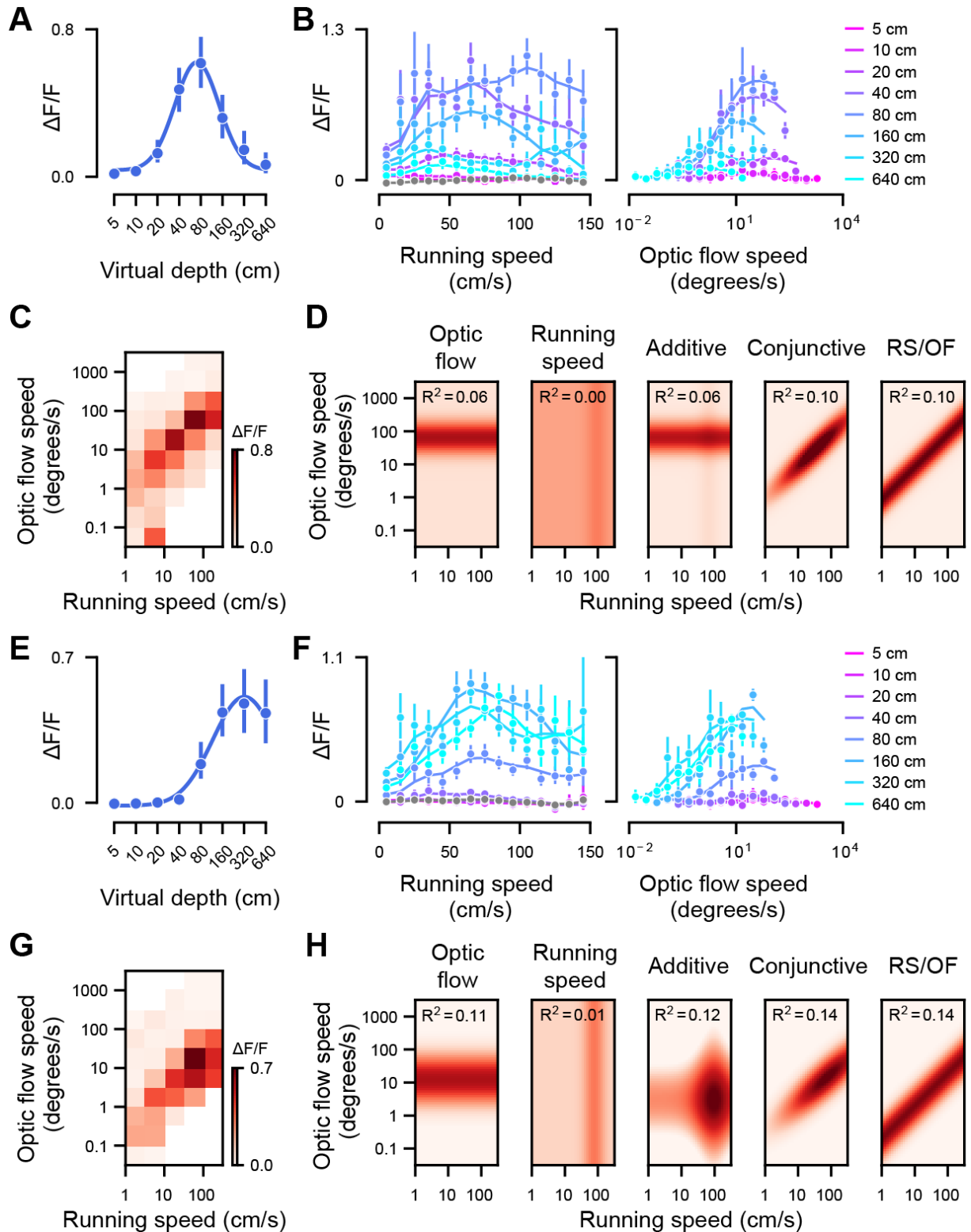


Figure 5.1: Activity of depth-selective neurons as a function of running and optic flow speeds. (*cont.*)

Figure 5.1: **Activity of depth-selective neurons as a function of running and optic flow speeds.** (A,E) Virtual depth tuning of two example depth-selective neurons. (B,F) Responses of the two example depth-selective neurons binned by running speed (B left, E left) or optic flow speed (B right, E right). Gray line (B left, E left) – running speed tuning during the inter-trial interval, when sphere stimuli were absent. As virtual depth equals to the ratio of running speed and optic flow speed, neuronal responses at a fixed running speed but different virtual depths (B left, E left) reflect the effect of optic flow speed. Similarly, neuronal responses at a fixed optic flow speed but different virtual depths reflect the influence of running speed (B right, E right). (C,G) Responses of the example neurons as a function of both running and optic flow speeds. (D,H) Running speed and optic flow tuning of the example neurons fitted with the pure optic flow, pure running speed, additive, conjunctive or ratio models (see methods). The heatmaps represent the fitted neuronal responses of each model as a function of running speed and optic flow speed, across the full range of speeds presented in the figure. RS – running speed, OF – optic flow speed.

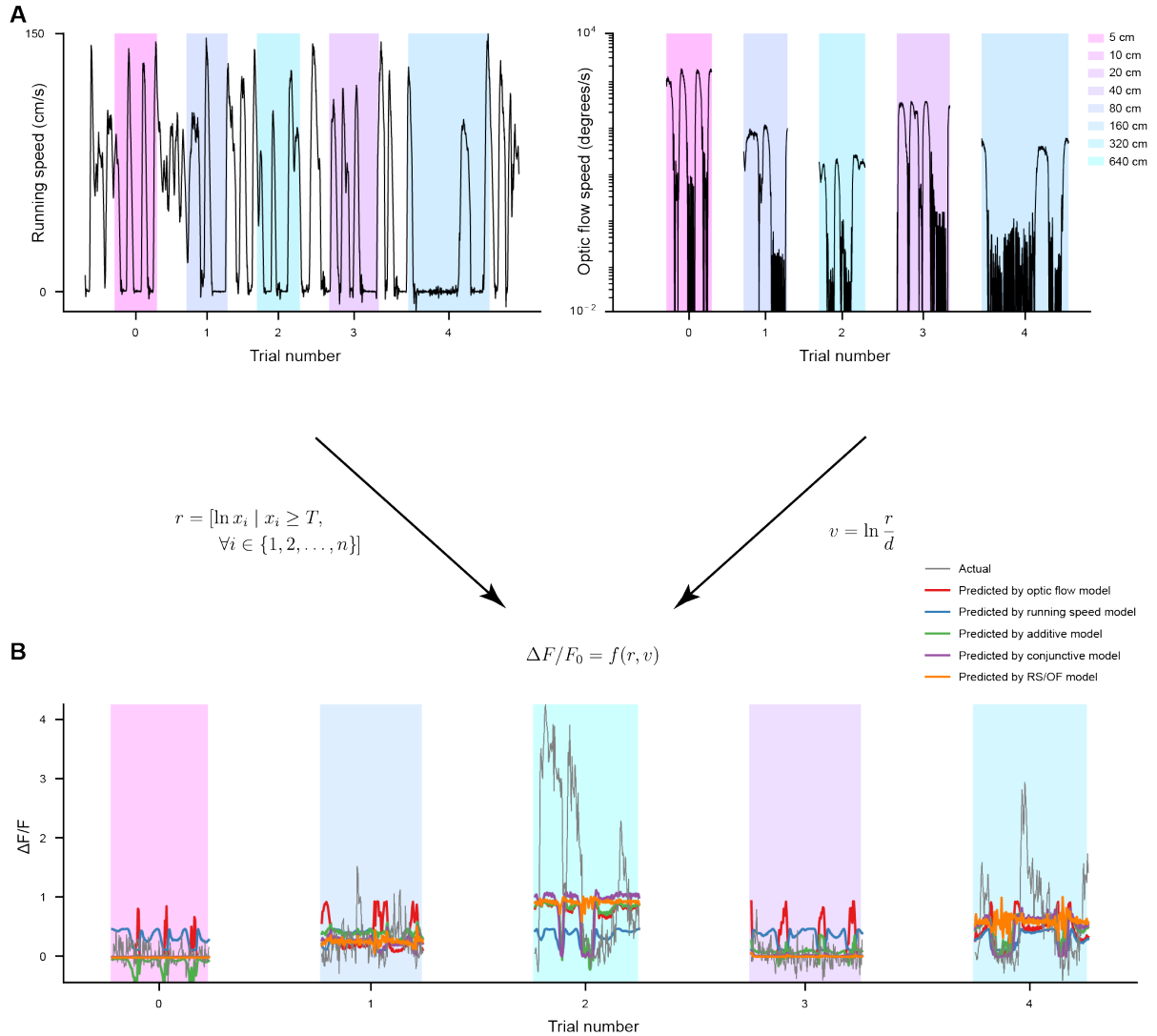


Figure 5.2: **Example running speed and optic flow speed traces and example fitted neuronal activity of the 5 models. Depth selectivity arises from conjunctive coding of optic flow speed and running speed.** (A) Example running speed trace (left) and optic flow trace (right) from 6 trials of an example session. Then, imaging frames with running speed (x) larger than a threshold (T , 1 cm/s) were taken, and running speed and the optic flow speed of the selected frames were log-transformed (right). The log-transformed running speed (r) and optic flow speed (v) were used to fit the 5 models (see Methods 2.4.11). (B) Actual and predicted $\Delta F/F_0$ of the example trials in (A) using the 5 models.

To compare the performance of models for each neuron, I used K-fold cross-validation to evaluate the goodness of fit (R^2) of each model for each neuron, using the test data that were excluded from parameter tuning. This helps to account for variations in model complexity. Then, I compared the proportion of neurons whose activity was best explained by each model in each session, applying different thresholds for depth selectivity and peak neuronal responses (Figure 5.3A-D). A larger peak response at the neuron's preferred virtual depth indicates a better signal-to-noise ratio, while a smaller p-value of the depth tuning fit signifies greater depth selectivity. Across thresholds, neither running speed nor optic flow speed alone could account for neuronal responses (Figure 5.3A-D), suggesting that the integration of optic flow and locomotion-related signals is required to produce depth selectivity. The conjunctive model and additive model consistently outperformed the optic flow model, the running speed speed model, or the RS/OF ratio model across recording sessions (Figure 5.3A-D). When looking at all depth-selective neurons, the conjunctive model performed similarly to the additive model (Figure 5.3A, $p = 0.0987$). However, when comparing model performance for highly depth-selective neurons with strong peak responses at their preferred virtual depth, the conjunctive model explained neuronal responses significantly better than the additive models across recording sessions (Figure 5.3B-D; $p < 0.0001$ for all thresholds). Taken together, depth-selective responses are best explained by a conjunctive coding of running speed and optic flow speed. Nevertheless, when comparing the observed neuronal activity with the predicted neuronal activity from the conjunctive model, a proportion of variance still seems unexplained by the model (Figure 5.2A-B). The underlying reason will be illustrated in the next chapter. In brief, beyond the conjunctive coding of running speed and optic flow speed, neuronal activity of depth-selective neurons is also influenced by the neurons' receptive fields. Since the current model lacks receptive field information, the predictions would be less well-explained compared to a model that incorporates the interactions between the stimuli and the neurons' receptive fields.

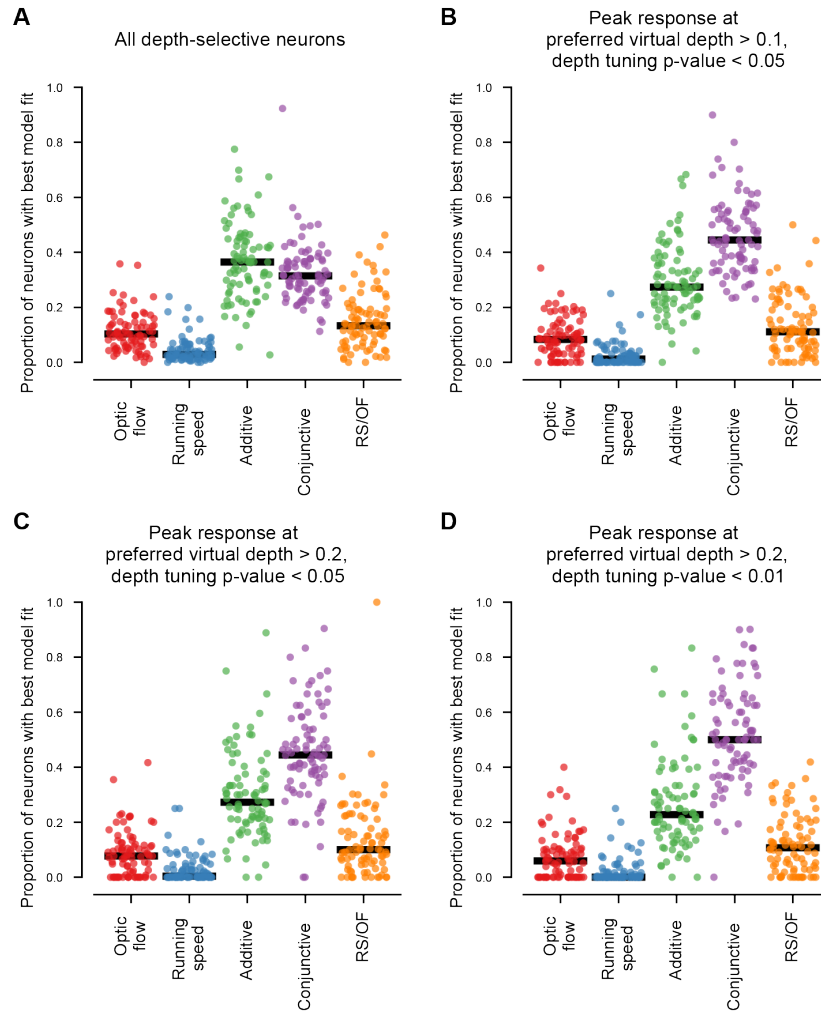


Figure 5.3: Depth selectivity arises from conjunctive coding of optic flow speed and running speed. (A-D) Proportion of depth-selective neurons best explained by the pure optic flow, pure running speed, additive, conjunctive and ratio models in each session, for neurons filtered with different criterion. (A) For all depth-selective neurons (N=31,013 from 85 sessions), the p-values for conjunctive model vs. other models are: vs. optic flow speed model: $p < 0.0001$, vs. running speed model: $p < 0.0001$, vs. additive model: $p = 0.0987$, vs. RS/OF model: $p < 0.0006$. (B) For depth-selective neurons with the peak responses at the preferred virtual depth > 0.1 (N=9,834 from 85 sessions), the p-values for conjunctive model vs. other models are: vs. optic flow speed model: $p < 0.0001$, vs. running speed model: $p < 0.0001$, vs. additive model: $p < 0.0001$, vs. RS/OF model: $p < 0.0001$. (C) For depth-selective neurons with the peak responses at the preferred virtual depth > 0.2 (N=7,337 from 85 sessions), the p-values for conjunctive model vs. other models are: vs. optic flow speed model: $p < 0.0001$, vs. running speed model: $p < 0.0001$, vs. additive model: $p < 0.0001$, vs. RS/OF model: $p < 0.0001$. (D) For depth-selective neurons with the peak responses at the preferred virtual depth > 0.2 and the p-value for depth-tuning fit < 0.01 (N=6,366 from 84 sessions, no neurons from 1 session pass the thresholds), the p-values for conjunctive model vs. other models are: vs. optic flow speed model: $p < 0.0001$, vs. running speed model: $p < 0.0001$, vs. additive model: $p < 0.0001$, vs. RS/OF model: $p < 0.0001$.

5.4 Preferred depth is determined by the ratio of preferred running speed and preferred optic flow speed

Given the observation above that the preferred conjunction of running speed and optic flow speed lies on the line corresponding to each neuron's preferred virtual depth, I explored the relationship between the preferred virtual depth of individual neurons and their preferred optic flow or running speed. To this end, I split the recordings for each neurons into two halves: half of the trials were used to estimate each neuron's preferred virtual depth, while the other half were used to estimate preferred optic flow and running speeds with the conjunctive model described above. Across the population, a strong negative correlation was observed between the optic flow speed preferences and virtual depth preferences (Figure 5.4A, $r = -0.728$, $p < 0.0001$). However, this correlation might arise from the limited range of optic flow speeds sampled in the VR stimuli, which was constrained by the possible running speeds of the mice (Figure 5.4B). Optic flow speed in the VR was calculated from the ratio of running speed and virtual depth. Given that the natural running speed of mice typically ranges from several cm/s to 2-3 m/s (Garland et al., 1995), the resulting optic flow speeds were restricted to a specific range, which might result in a natural negative correlation with virtual depth (Figure 5.4B). This relationship is demonstrated by the overlap between the preferred optic flow speeds of depth-selective neurons and the range of optic flow speeds sampled (Figure 5.4B). It remains uncertain whether the observed distribution of optic flow speed preferences would persist if a broader range of running speeds were sampled.

The running speed preferences and virtual depth preferences of all depth-selective neurons did not seem to have an apparent correlation (Figure 5.4C, $r = 0.013$, $p = 0.715$). However, when considering neurons tuned to low (<10 degrees/s), intermediate (10-100 degrees/s) or high (100-1000 degrees/s) optic flow speeds separately, preferred running speed displayed a significant positive correlation with preferred virtual depth (Figure 5.4C, low optic flow speeds, $r = 0.617$, $p < 0.0001$; intermediate optic flow speeds, $r = 0.568$, $p < 0.0001$; high optic flow speeds, $r = 0.489$, $p < 0.0001$). This pattern reflects the Simpson's paradox, where a statistical trend is evident within subgroups of data but disappears when the groups are combined.

Overall, selectivity for both optic flow speed and running speed contribute to the virtual depth preferences of individual neurons. This population-level trend also mirrors the principles of motion parallax, that far depths generate slower optic flow speed when controlling for running speed, and faster running is required to generate the same optic flow speed for far depths.

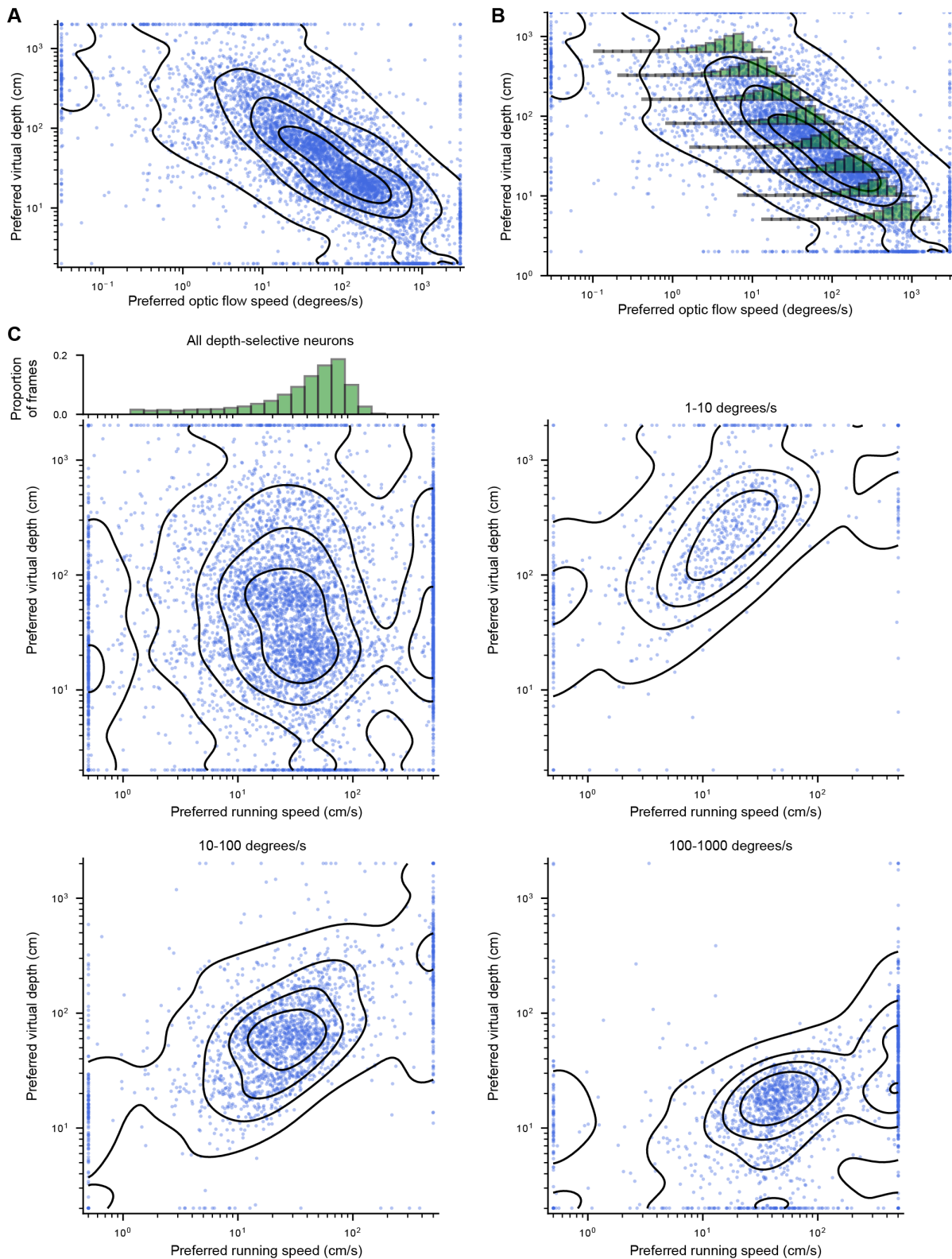


Figure 5.4: The relationship between preferred depth, preferred running speed and preferred optic flow speed. (cont.)

Figure 5.4: **(A)** Preferred virtual depth as a function of preferred optic flow speeds estimated using hold-out data not used in calculating preferred virtual depths ($r = -0.728$, $p < 0.0001$). **(B)** Preferred virtual depth as a function of preferred optic flow speeds in **(A)** overlapped with histograms of possible optic flow distribution across virtual depths calculated from the distribution of running speed across all sessions. The y-axis of the histograms is linear and does not corresponds to the y-axis of preferred virtual depth. **(C)** Preferred virtual depth as a function of preferred running speeds of all neurons estimated using hold-out data not used in calculating preferred virtual depths (top left, $r = 0.013$, $p = 0.715$), and for neurons with preferred optic flow speeds between 110 degrees/s (top right, $r = 0.617$, $p < 0.0001$), 10100 degrees/s (bottom left, $r = 0.568$, $p < 0.0001$), and 1001000 degrees/s (bottom right, $r = 0.489$, $p < 0.0001$). Histogram represents the distribution of running speed across all sessions (top left).

Next, I examined the relationship between the ratio of preferred running speed to preferred optic flow speed and virtual depth preferences of individual neurons. Consistent with the observed relationship in Figure 5.1C and F, I found that the ratio of the preferred running and optic flow speeds of individual V1 neurons could predict their preferred virtual depth (Figure 5.5A, $r = 0.794$, $p < 0.0001$). Most neurons clustered around the expected gradient, where preferred virtual depths equal to the ratio of the preferred running and optic flow speeds (Figure 5.5A). Nevertheless, since these two variables were estimated from independent trials and given the large sample size, it is reasonable the regression slope did not exactly equal one (Huber regression, gradient = 0.533).

This relationship can be further illustrated by visualising running speed and optic flow speed preferences of neurons preferring various virtual depths in logarithmic space (Figure 5.5B). As different virtual depths are associated with different ratios of running and optic flow speeds, expected log-transformed optic flow speeds for each virtual depth lie along a line as a function of log-transformed running speed with a different intercept in logarithmic space (Figure 5.5B inset). Depth-selective responses echoed this relationship, where neurons tuned to the same virtual depth lie along a line in the log-transformed preferred running speed / preferred optic flow speed space. In other words, neurons tuned to the same virtual depth preferred different combinations of running and optic flow speeds with a constant ratio. As a consequence, neurons tuned to similar virtual depth have a wide range of running speed preferences, enabling neuronal populations in V1 to encode depth regardless of animal's locomotion speed.

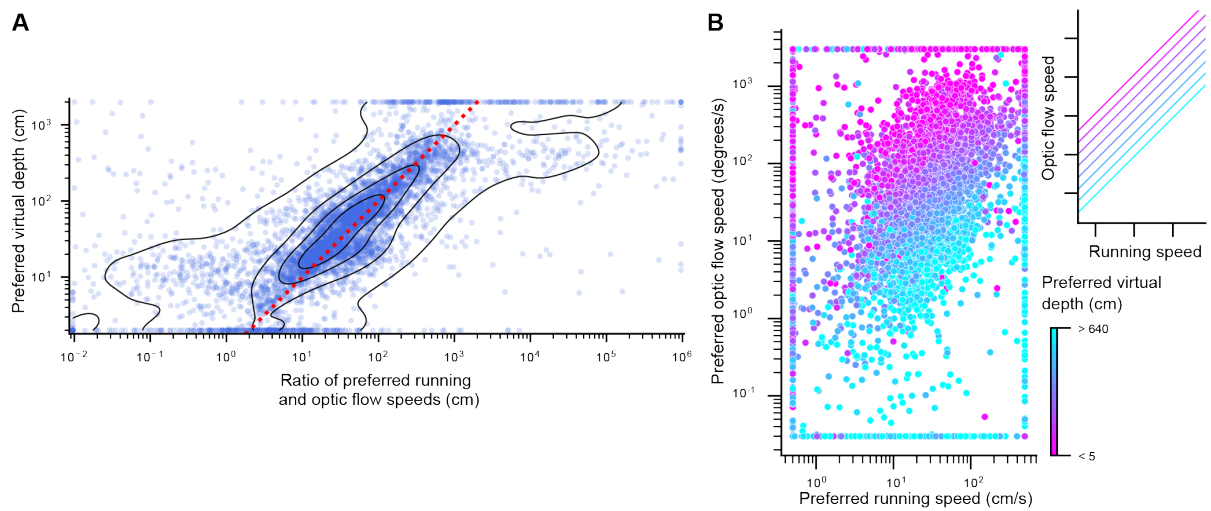


Figure 5.5: **Preferred depth is determined by the ratio of preferred running speed and preferred optic flow speed.** (A) Preferred virtual depths as a function of the ratio between preferred running and optic flow speeds estimated using hold-out data not used in calculation of preferred virtual depths ($N = 5,737$ neurons). Contours – kernel density estimate. Red dotted line – when preferred virtual depths equal to the ratio between preferred running and optic flow speeds estimated by hold-out data. (B) Preferred optic flow speed and running speed of neurons estimated using hold-out data color coded by their preferred virtual depth ($N = 5,737$ neurons). Inset – optic flow speed as a function of running speed across virtual depths.

5.5 Closed loop coupling between locomotion and optic flow enables accurate representation of depth

Next, I explored whether closed loop coupling between locomotion and optic flow feedback was necessary for the integration of two signals to produce depth-selective responses. To this end, I broke the coupling between locomotion and optic flow in an open-loop manner, by manipulating the optic flow feedback using a previously recorded running trajectory (Figure 5.6A-B). In this way, optic flow speed in the VR environment was determined by this previously recorded running trajectory instead of the current running speed of the mouse, which broke the closed loop coupling (Figure 5.6C-D) while maintaining the overall statistics of the visual stimuli.

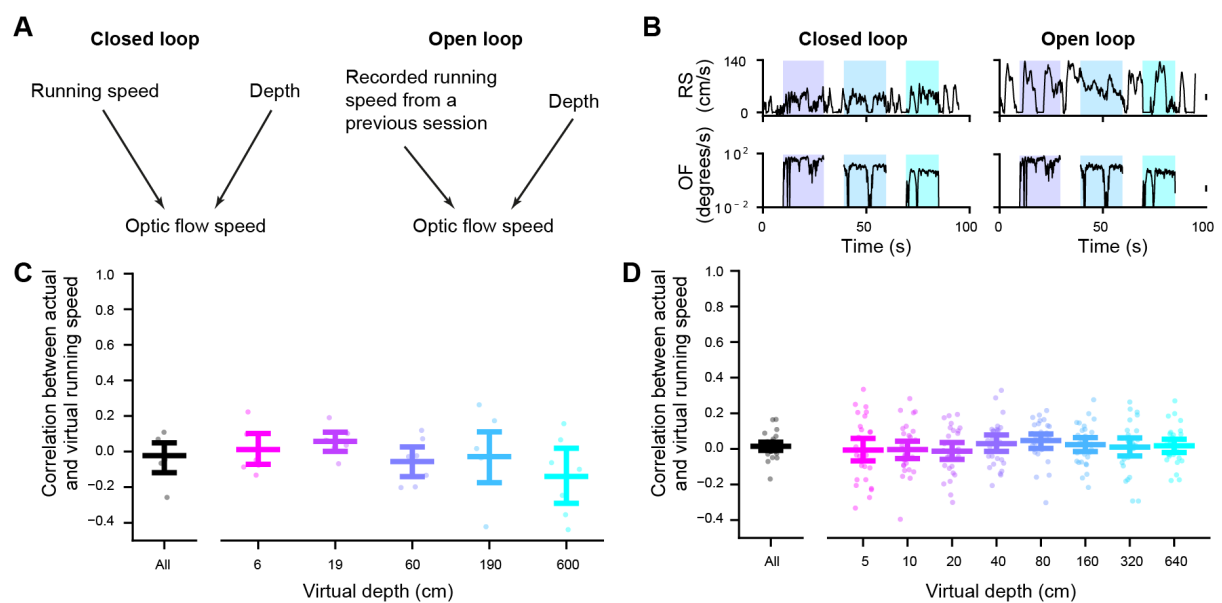


Figure 5.6: **Open loop stimuli breaks the coupling between locomotion and optic flow feedback.** (A-B) Relationship between running speed and optic flow speed on closed loop and open loop trials (A) and example traces (B). Shaded areas color-coded by depth indicate stimulus presentation. (C-D) Correlation coefficient between the actual and virtual running speeds during open loop recordings with 5 (C) or 8 (D) virtual depths for individual imaging sessions.

I then recorded neuronal responses during blocks of closed loop and open loop trials and analysed whether the running speed and optic flow speed preferences changed when the closed loop coupling was broken. For individual neurons, preferred optic flow speed and running speed remained similar when the coupling between locomotion and optic flow feedback was disrupted by the open loop stimuli (Figure 5.7A-D). Across all depth-selective neurons, optic flow speed and running speed preferences were highly correlated between closed loop and open loop conditions (Figure 5.7E-F, running speed $r = 0.578$, $p < 0.0001$, optic flow speed $r = 0.700$, $p < 0.0001$). The peak response magnitude at the combination of preferred running and optic flow speeds was maintained during open loop recordings (Figure 5.7G, $p = 0.681$). These results suggest that the conjunctive coding of locomotion and optic flow speeds does not depend on their closed loop coupling.

Photobleaching during imaging could reduce the fluorescence responses overtime if some trials were recorded much later than the others (Patterson & Piston, 2000). Since the closed loop recordings usually started 1-2 hours before the open loop recordings, the unchanged peak response amplitude could be a result of photobleaching. To test this effect, I compared the peak response amplitudes of neurons recorded in a subset of sessions, where there were closed loop trials both before and after open loop trials. Even if the peak response amplitudes of neurons seemed slightly higher in closed loop recordings preceding an open loop recording, they remained statistically similar regardless of the order of closed loop and open loop trials (Figure 5.7H-I; ratio of response amplitudes in closed loop before open loop trials vs. open loop trials: median 1.394, $p = 0.0774$; closed loop after open loop vs. open loop trials: median 0.961, $p = 0.783$). Therefore, although slight photobleaching could be happening during the recording, it does not significantly affect the peak fluorescence amplitude of depth-selective neurons.

I then examined whether closed loop coupling enhanced the accuracy of depth representation at the population level. To test whether virtual depth was more accurately represented by population activity in closed loop condition, I trained a linear support vector machine (SVM) classifier to decode virtual depth from activity of neuronal populations recorded in closed loop recordings and open loop recordings respectively. I split imaging trials into training, validation and test sets with 5-fold cross validation such that each trial was included in one test set. Then, I trained the classifier on individual imaging frames from trials in the training set, optimised the hyperparameter using the validation set, and finally evaluated the performance of the optimised classifier using the test set. Confusion matrices of decoded virtual depth demonstrated that the classifier could decode the virtual depth of individual frames from population activity in V1 in closed loop recordings, but performed poorly in open loop recordings (Figure 5.8A-D). The accuracy of the classifier was significantly higher in closed loop recordings compared to open loop recordings (Figure 5.8E). These results indicate that closed loop coupling between locomotion and optic flow is required for accurate representation of depth by neuronal populations in V1, and suggest that locomotion-related signals which give rise

5.5. Closed loop coupling between locomotion and optic flow enables accurate representation of depth

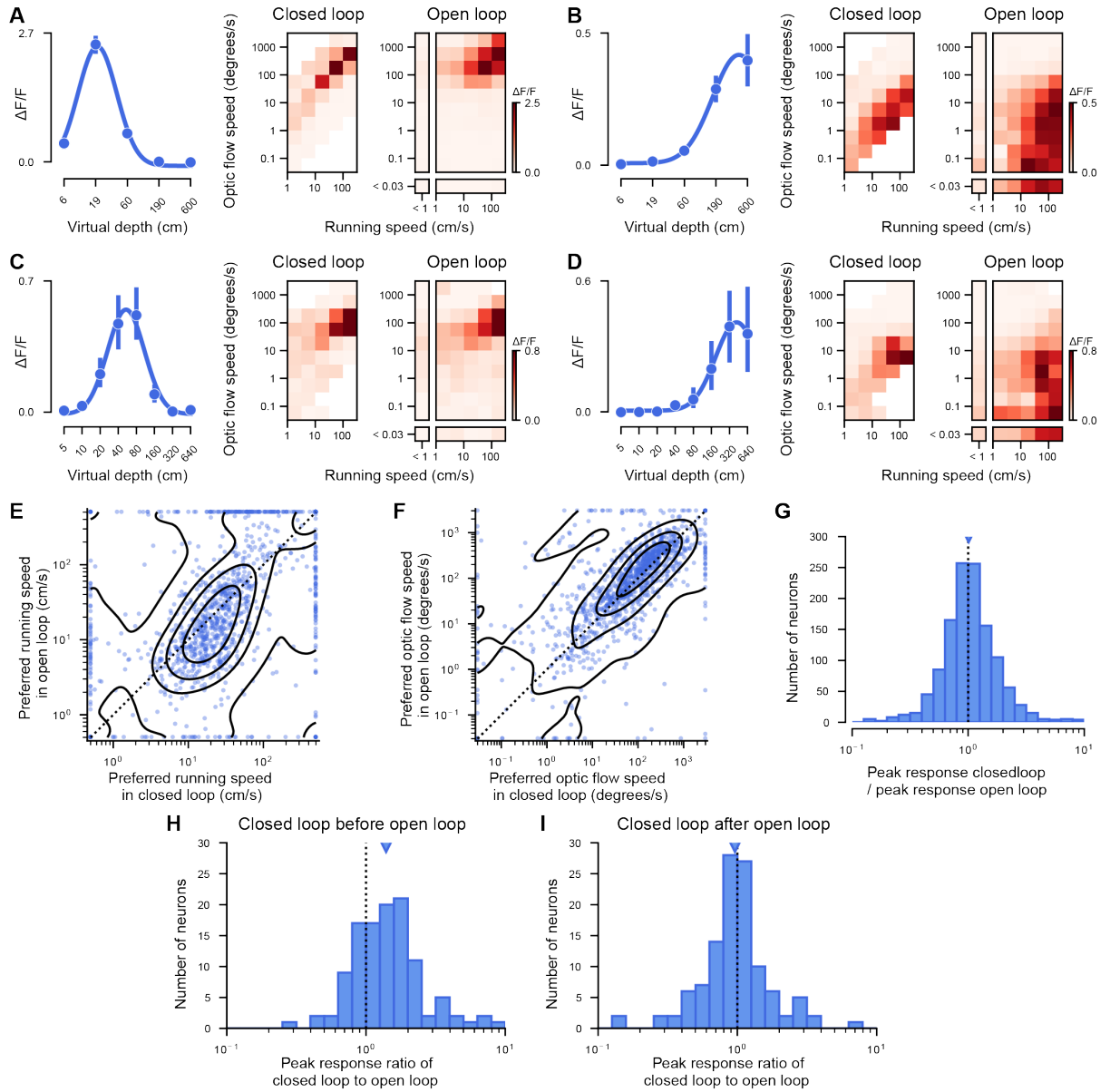


Figure 5.7: Conjunctive coding of locomotion and optic flow speeds does not rely on their closed loop coupling. (A-D) Depth tuning (left) and the neuronal responses as a function of both running and optic flow speeds in closed loop (middle) and open loop (right) in 4 example neurons. (E-F) Scatter plot of preferred running speed (E), preferred optic flow speed (F) in open loop and closed loop for depth-selective neurons (N = 1,234 neurons from 34 sessions). Contours – kernel density estimates. (G) Histogram of ratios of peak response amplitude in closed loop over open loop for neurons in J-K. (H-I) Histogram of ratios of peak response amplitude in closed loop over open loop for closed loop trials before (H) or after (I) open loop trials (N = 116 neurons from 10 sessions in 4 mice).

to the optic flow feedback may contribute to disambiguating the depth of visual stimuli.

Since closed loop recordings usually started 1-2 hours before open loop recordings in most sessions, I also assessed whether photobleaching during imaging could affect decoding accuracy. I compared the decoding accuracy between closed loop and open loop recordings in a subset of sessions where closed-loop recordings occurred both before and after the open

loop trials. Due to insufficient trial numbers for each depth in some recordings, I focused on 5 sessions where the number of trials allowed for 4-fold cross-validation (Figure 5.8F). Although statistical significance was not reached due to the limited sample size (Figure 5.8F, closed loop before open loop vs. open loop: Wilcoxon signed-rank test, $p = 0.4375$; closed loop after open loop vs. open loop: Wilcoxon signed-rank test, $p = 0.0625$), the decoding accuracy of the closed loop recordings after an open loop recording was, on average, higher than the decoding accuracy of the open loop recordings within the same session (Figure 5.8F). This suggests that the reduced classification accuracy in open loop recordings is unlikely to be caused by photobleaching. The relatively low decoding accuracy of the closed loop recordings preceding an open loop recording is likely due to the limited amount of data (typically comprising fewer than 5 trials per virtual depth).

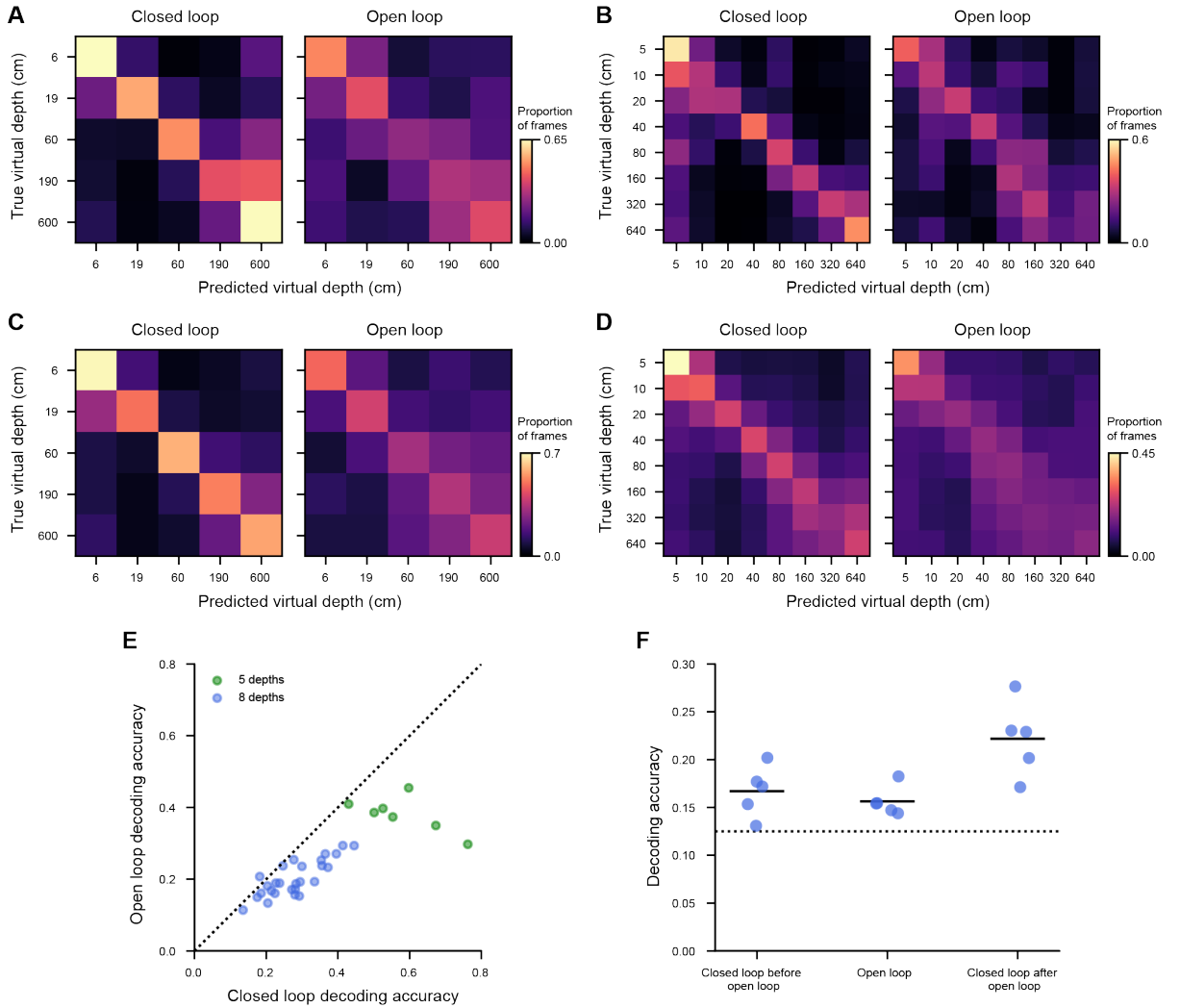


Figure 5.8: Closed loop coupling between locomotion and optic flow enables accurate representation of depth by neuronal populations in V1. (A) Confusion matrices of decoded virtual depth in closed loop and open loop recordings of an example session with 5 virtual depths. (B) Confusion matrices of decoded virtual depth in closed loop and open loop recordings of an example session with 8 virtual depths. (C) Confusion matrices of decoded virtual depth in closed loop and open loop recordings with 5 virtual depths (N = 7 sessions). (D) Confusion matrices of decoded virtual depth in closed loop and open loop recordings with 8 virtual depths (N = 27 sessions). (E) Accuracy of decoding virtual depth from population activity of depth-selective neurons in closed loop and open loop recordings (N = 34 sessions). (F) Accuracy of decoding virtual depth from population activity of depth-selective neurons in closed loop recordings preceding an open loop recording, open loop recordings, and closed loop recordings after an open loop recording (N = 5 sessions).

5.6 Speed preferences of V1 neurons are consistent across conditions of passive viewing and and self-generated visual flow

Traditionally, the speed tuning of V1 neurons has been characterised using passive viewing of two-dimensional stimuli, such as bars or drifting gratings. (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). In the VR setup, optic flow was generated by active self-motion within a three-dimensional environment. It was unclear whether the optic flow speed preferences of V1 neurons mapped in three-dimensional VR environments during active locomotion of mice correspond to the speed tuning mapped with canonical two-dimensional stimuli during passive viewing. To test this question, I recorded neuronal activity with both the three-dimensional VR stimuli and two-dimensional drifting gratings in a subset of recordings. The VR stimuli were updated in closed loop according to the animal's running speed, and the drifting gratings were uncoupled with the animal's locomotion. Gratings were presented in combinations of 6 spatial frequencies, 6 temporal frequencies, and 8 orientations in a pseudorandom order, and each set of gratings were presented for 5 times throughout the recording sessions. A direction of 0° represented vertical gratings drifting forward, and a direction of 90° represented horizontal gratings drifting downward.

First, I found that neurons exhibiting strong responses to drifting gratings could be either depth-selective (Figure 5.9A-D) or non-depth-selective (Figure 5.9E-H, Figure 5.10A). To characterise the spatiotemporal tuning of neurons, I fitted the responses of all neurons as a function of spatial frequency, temporal frequency and grating direction. The model combined a spatiotemporal tuning component as a bivariate Gaussian function and a direction tuning component following a von Mises distribution (Figure 5.9C-D, G-H, see Methods 2.4.11). Then I defined the preferred spatial frequency and preferred temporal frequency as the combination of spatiotemporal frequencies that gave rise to the peak response amplitude. The majority (58%) of neurons displayed strong spatiotemporal tuning mapped by drifting gratings (Figure 5.10A), including both populations of depth-selective and non-depth-selective neurons (Figure 5.10A).

Then, I quantified the relationship between the preferred optic flow speed mapped by three-dimensional VR stimuli and the preferred speed of neurons mapped by grating stimuli. To this end, I selected neurons that have both strong spatiotemporal frequency tuning (R^2 of the spatiotemporal tuning fit > 0.05) and strong conjunctive tuning to running speed and optic flow speed (cross-validated R^2 of the conjunctive model fit > 0.02). Although the preferred speed of most V1 neurons changes with spatial frequency (LeDue et al., 2012), for simplicity, preferred speed of a neuron here was defined as the ratio of its preferred temporal frequency to its preferred spatial frequency. I found that optic flow speed preferences were positively correlated with the speed preferences mapped by gratings (Figure 5.10B, $r = 0.384, p = 0.0109$, p-value was calculated directly from

Spearman's correlation). This result demonstrates some consistency between the speed tuning of V1 neurons mapped by passive two-dimensional stimulation and self-generated visual flow, although additional samples are required to strengthen this conclusion.

It is worth noting that optic flow speed preferences mapped by the VR stimuli only included two directions (parallel to the animal's direction of running), whereas the speed preferences mapped by gratings include 8 directions of motion. To explore the distribution of preferred directions of depth-selective neurons mapped by passive viewing, I first visualised the distribution of the direction selectivity index (DSI) across all depth-selective neurons with strong spatiotemporal frequency tuning ($N = 477$ out of 3,012 neurons, Figure 5.10C). Among depth-selective neurons, many did not display strong direction selectivity, with half of the neurons having a DSI lower than 0.5 (Figure 5.10D). Many depth-selective neurons displayed preferences to horizontal gratings (Figure 5.10E). Among depth-selective neurons with strong direction selectivity ($DSI > 0.6$, $N = 193$ neurons), neurons that preferred horizontal gratings drifting upward were enriched (Figure 5.10F). The orientation and direction preferences mapped by gratings are consistent with the orientations of RFs mapped by VR stimuli typically found in depth-selective neurons (Figure 6.2A), which will be demonstrated in the next chapter.

Interestingly, the preferred speed mapped by gratings showed a negative correlation with preferred virtual depth (Figure 5.10C, $r = -0.315$, $p = 0.0396$, p-value was calculated directly from Spearman's correlation). Despite being significant within a small sample, this is a much weaker correlation compared to the correlation between the preferred virtual depth and the ratio between optic flow and running speed preferences. Therefore, visual motion speed tuning during passive viewing is a less reliable predictor of depth tuning compared to the combination of running speed and self-generated optic flow speed signals.

5.6. Speed preferences of V1 neurons are consistent across conditions of passive viewing and self-generated visual flow

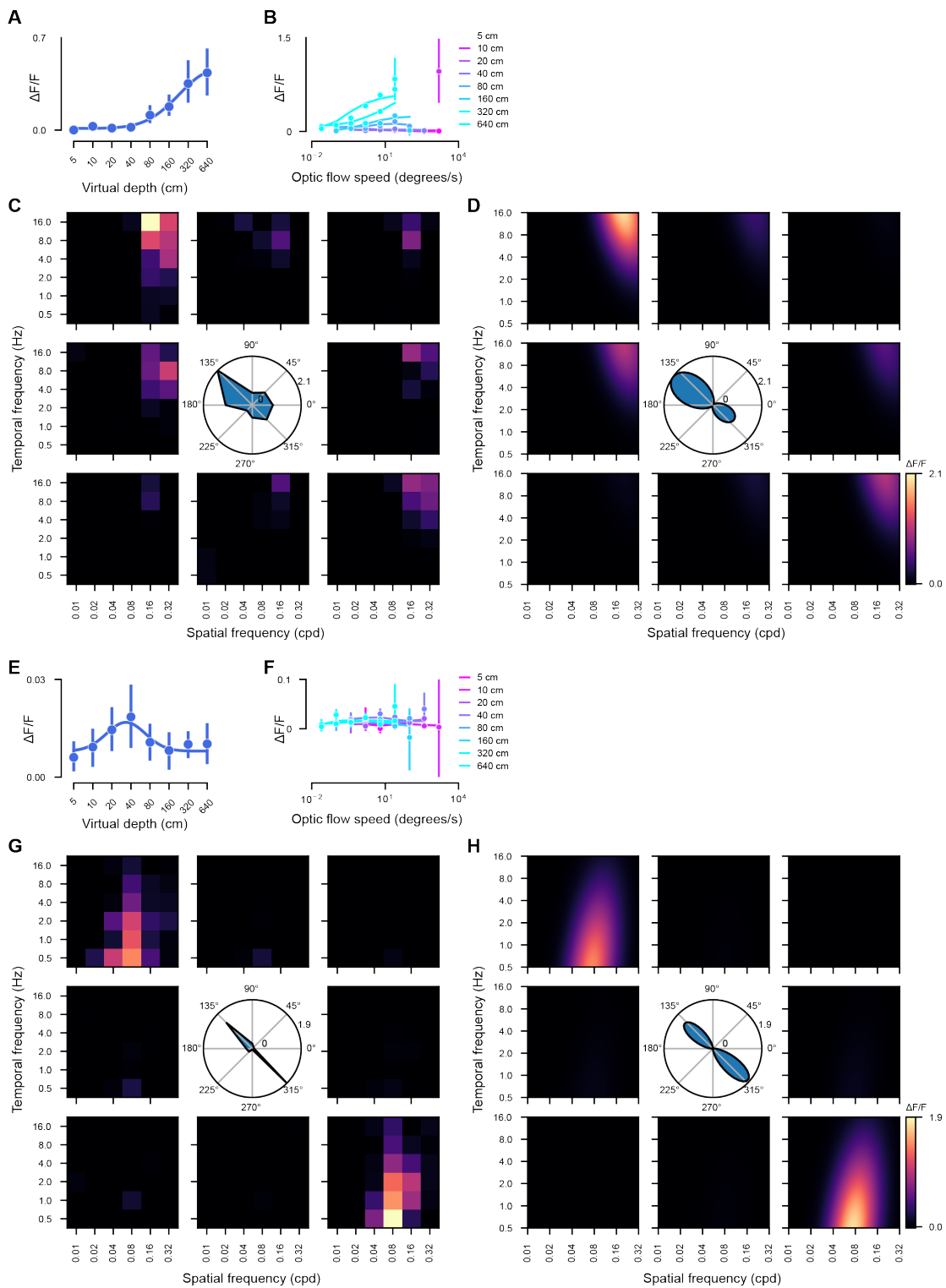


Figure 5.9: **Spatiotemporal tuning of example V1 neurons.** (A-D) An example depth-selective neuron, its depth tuning curve (A), optic flow speed tuning as a function of virtual depth (B), responses to 8 orientations of drifting gratings binned by spatial and temporal frequencies (C), and the fitted responses as a function of spatial and temporal frequencies at 8 orientations of drifting gratings (D). (E-H) An example non-depth-selective neuron, its depth tuning curve (E), optic flow speed tuning (F), and spatiotemporal tuning (G-H).

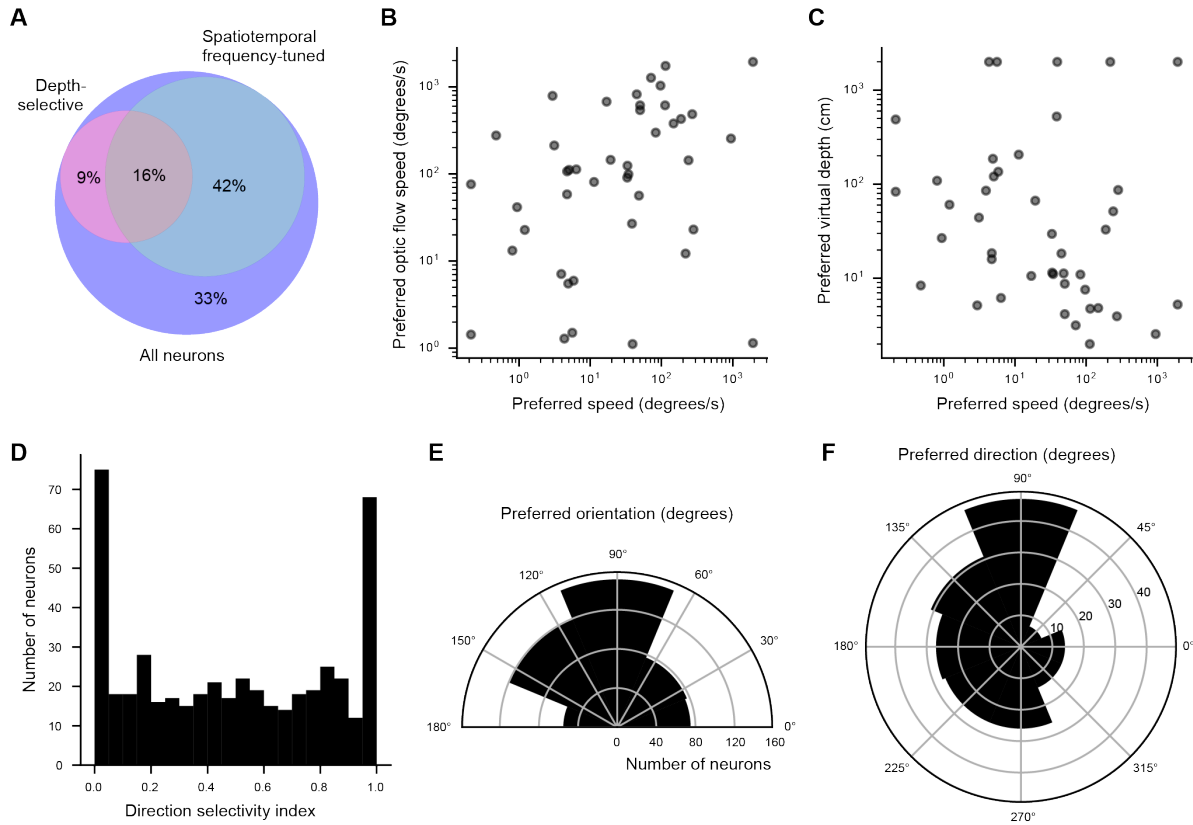


Figure 5.10: **Speed preferences mapped by three-dimensional stimuli and two-dimensional gratings.** (A) Venn diagram of depth-selective neurons and neurons with strong spatiotemporal frequency tuning. (B) Scatter plot of preferred optic flow speed against preferred speed mapped by gratings of neurons significantly tuned to the conjunctive model and spatiotemporal frequency (43 out of 3,012 neurons from 6 sessions from 2 mice). (C) Scatter plot of preferred virtual depth against preferred speed mapped by gratings of neurons in (B). (D) Histogram of the direction selectivity index of all depth-selective neurons that have strong spatiotemporal frequency tuning (N = 477 out of 3,012 neurons). (E) Histogram of the preferred orientation of all depth-selective neurons that have strong spatiotemporal frequency tuning (N = 477 out of 3,012 neurons). 0° and 180° represent vertical gratings, and 90° represents horizontal gratings. (F) Histogram of the preferred orientation of all depth-selective neurons that have strong spatiotemporal frequency tuning and strong direction selectivity (N = 193 out of 3,012 neurons). 0° or 180° represents vertical gratings drifting forward or backward. 90° or 270° represents horizontal gratings drifting downward or upward.

5.7 Discussion

5.7.1 Locomotion-related modulation supports depth representation in V1

Locomotion-related modulation of neuronal responses in mouse V1 has been extensively characterised. Locomotion elevates the amplitude of visually-evoked responses (Ayaz et al., 2013; Keller et al., 2012; Niell & Stryker, 2010), regulates tuning to certain visual features (Ayaz et al., 2013; Mineault et al., 2016), and improves the accuracy of visual representation among V1 populations (Dadarlat & Stryker, 2017; Erisken et al., 2014; McBride et al., 2019). However, the functional significance of locomotion-related modulation of V1 neurons in visual processing remains heavily debated (Christensen & Pillow, 2022; Dadarlat & Stryker, 2017; Keller & Mrsic-Flogel, 2018; Keller et al., 2012; Muzzu & Saleem, 2021; Niell & Stryker, 2010). Here I propose that locomotion-related modulation allows L2/3 neurons in V1 to infer depth of visual cues in the environment – a crucial computation that needs to be conducted by the visual system to support animals’ daily life. By analysing depth-selective responses of V1 neurons as a function of running speed and optic flow speed, I show that depth selectivity arises from the integration of these two signals. Depth selectivity is generated from a conjunctive tuning of running and optic flow speed. For individual neurons, their depth preferences are determined by the ratio of preferred running and optic flow speed. As a population, neurons preferring similar virtual depth encompasses a wide range of running speed preferences, enabling V1 population to represent a variety of depths when animals are running at any speed.

It has been long proposed that locomotion can adjust the gain of visual responses in V1 excitatory neurons, amplifying visual responses during the state of running (Erisken et al., 2014; Niell & Stryker, 2010). While some studies proposed that the state of locomotion induces a generic modulation of cortical state which boosts the gain of sensory responses across V1 (Bennett et al., 2013; Harris & Thiele, 2011; Niell & Stryker, 2010), others have found that the effect of running on visual responses depends on the running speed tuning of individual neurons (Christensen & Pillow, 2022; Keller et al., 2012; Saleem et al., 2013). V1 consists of neurons that monotonically increase or decrease their activity as a function of running speed, and neurons that are non-monotonically tuned to running speed (Christensen & Pillow, 2022; Saleem et al., 2013). Consistent with previous literature, I found that V1 neuronal responses arise from the integration of visual motion speed and running speed when mice experience optic flow generated by their own locomotion in a 3D VR environment. Individual neurons have distinct optic flow speed tuning and running speed tuning. The depth-selective responses I characterised here is achieved by multiplicative gain modulation – locomotion-related signals modulate the gain of optic flow responses in individual neurons, with the greatest responses occurring at the conjunction of a neuron’s preferred running speed and optic flow speed. The conjunctive coding of running speed and visual motion speed observed in my study differs from the findings of Saleem et al.

(2013), who reported that V1 neuronal responses could be explained by a linear summation of running speed and optic flow speed signals when mice were navigating in a VR corridor with grating and plaid patterns. This conjunctive integration may not have been evident in Saleem et al. (2013) due to the narrower range of running speeds examined in their study. Specifically, the running speed of mice in their experiments reached up to 30 cm/s, whereas the preferred running speeds of most neurons characterised in my studies exceed 10 - 30 cm/s.

Previous studies have shown that gain modulation can support coordinate transformation (Andersen et al., 1985; Lu et al., 2022; Lyu et al., 2022; Pouget & Sejnowski, 1997; Zipser & Andersen, 1988). For example, it has been proposed that posterior parietal neurons in monkeys can support the transformation from retinal to body-centric coordinates (Andersen et al., 1985; Pouget & Sejnowski, 1997; Zipser & Andersen, 1988). It is achieved by jointly encoding different combinations of retinotopic inputs and eye positions through gain modulation. Then, the activity of joint-coding neurons can be converged to produce egocentric representation of the visual scene invariant to eye positions. Another example is the central complex in the brain of *Drosophila melanogaster*, which can transform inputs from egocentric to allocentric coordinates in a similar manner (Lu et al., 2022; Lyu et al., 2022). PFN neurons receive inputs from EPG terminals in the protocerebral bridge which conveys information on world-centric heading direction, and project to h Δ B neurons in the fan-shaped body which outputs the world-centric travelling direction of the fly. To compute allocentric travelling directions and guide navigation, PFN neurons integrate different combinations of world-centric heading directions and body-centric translational velocity signals in a multiplicatively manner. Then, PFN neurons preferring similar world-centred travelling directions are converged into h Δ B neurons. In this way, body-centric translational vectors are transformed into world-centric travelling directions regardless of body orientations. In our case, I propose that L2/3 neurons in V1 can integrate optic flow and locomotion inputs and generate sets of basis functions tuned to various combinations of running and optic flow speed through gain modulation. Then downstream areas, such as V1 L5/6 or HVAs, can integrate inputs from L2/3 neurons with similar depth preferences, leading to depth representation by neurons that are tuned to depth independent of running or optic flow speed (Figure 5.11).

5.7.2 Predictive coding

Another popular hypothesis on the visual functions of locomotion-related modulation is associated with predictive processing (Jordan & Keller, 2020; Keller & Mrsic-Flogel, 2018; Keller et al., 2012; Muzzu & Saleem, 2021; Zmarz & Keller, 2016), which suggests that neuronal activity is driven by the comparison between bottom-up sensory inputs and top-down prediction based on an internal model of the world (Rao & Ballard, 1999). In V1, it has been proposed that L2/3 neurons may signal visuomotor prediction errors by computing the discrepancies between actual visual flow and predicted visual feedback

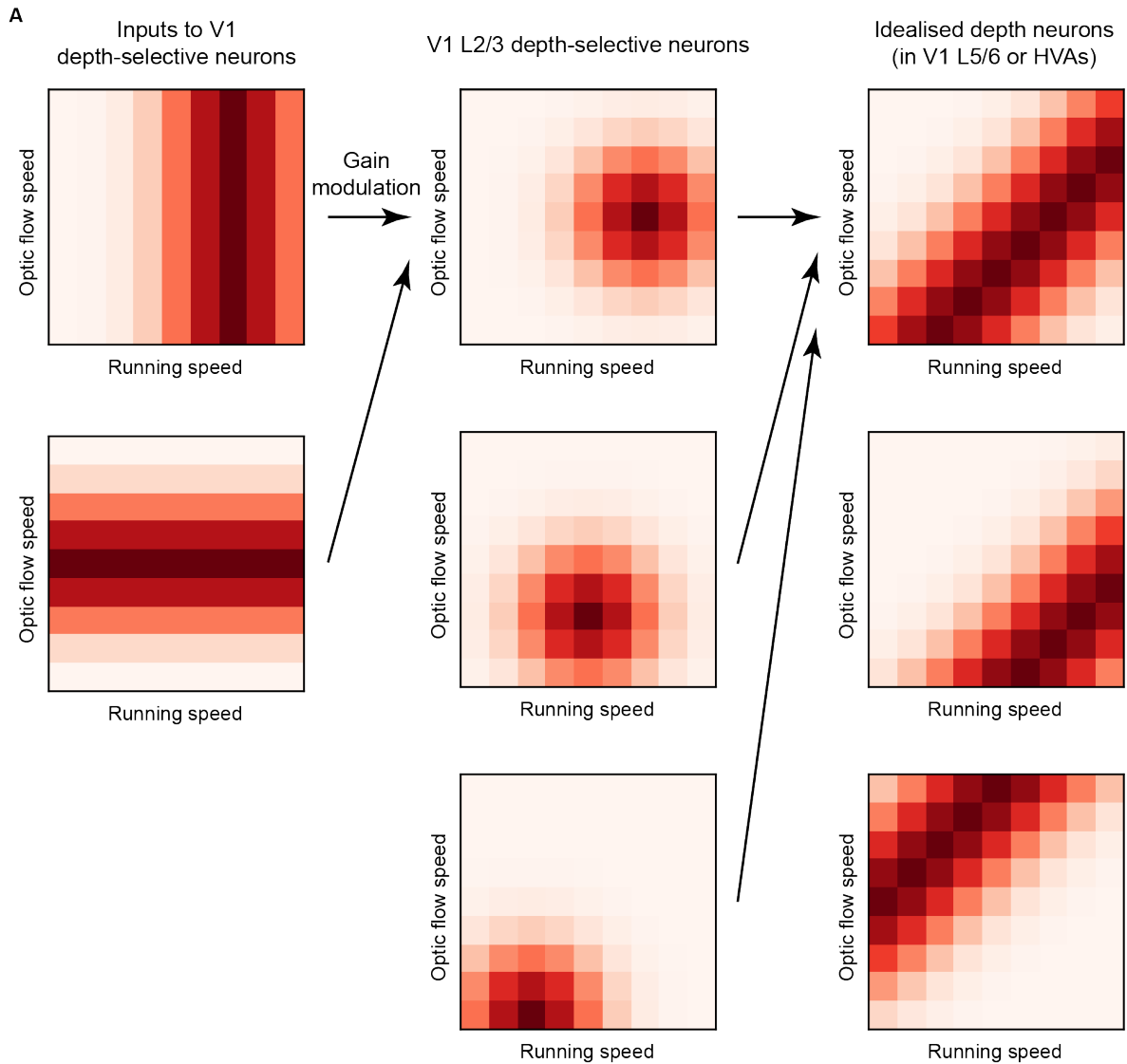


Figure 5.11: **Model of depth representation in the visual cortex.** (A) Schema of depth representation in the visual cortex. Depth-selective neurons in the L2/3 of V1 receives running speed and optic flow speed inputs (left) and integrate the two signals via gain modulation. This leads to neurons preferring different combinations of running speed and optic flow speed (middle). Then, the outputs of L2/3 neurons preferring similar depth are converged in deeper layers in V1 or HVAs to form neurons that are tuned to depth invariant to running or optic flow speed (right).

based on locomotion-related input (Keller & Mrsic-Flogel, 2018). A subset of L2/3 V1 neurons were found responsive to a mismatch between running speed and visual flow speed created by a sudden halt in visual flow (mismatch neurons), whereas passive viewing of the playback of visual flow halts did not elicit responses in these neurons (Jordan & Keller, 2020; Keller et al., 2012; Zmarz & Keller, 2016). However, my findings do not support the hypothesis that V1 neuronal responses are driven by visuomotor prediction errors. First, as optic flow speed is determined by both depth and running speed, optic flow speed cannot be predicted without the information on depth. Second, depth-selective neurons were active throughout the trials of their preferred virtual depth despite no prediction

errors were created under the closed loop condition. Third, the conjunctive coding of visual motion and running speed does not rely on closed loop coupling of locomotion and optic flow feedback. Running speed and optic flow speed preferences as well as the peak response amplitude of individual neurons were maintained even when prediction errors were introduced in the open loop condition. Therefore, depth-selective responses shown in my experiments do not reflect mismatch signals.

Could the mismatch responses observed in previous studies (Attinger et al., 2017; Jordan & Keller, 2020; Keller et al., 2012; Leinweber et al., 2017; Muzzu & Saleem, 2021; Zmarz & Keller, 2016) be interpreted through the lens of the depth representation model? Neurons selective for negative prediction errors (slower visual flow than predicted or visual flow halts during locomotion) may actually be responding to a lower ratio between visual flow speed and running speed, corresponding to far depths. In contrast, neurons selective for positive prediction errors (faster visual flow than predicted) may be responding to near depths. On the other hand, predictive coding and depth representation could co-exist as functions of V1 neurons.

Another possibility is that prediction error representation of visuomotor mismatch may still exist in V1 and the rest of the visual cortex. Prediction errors signals and depth could be represented by independent populations in V1. While depth selectivity is widespread in V1, mismatch neurons identified in previous studies only occupy a small proportion of V1 neurons (Jordan & Keller, 2020; Keller et al., 2012; Zmarz & Keller, 2016). Depth-selective responses in V1 can also form the basis for predictive coding in downstream areas. Depth information can be used to refine the prediction of optic flow speed, enabling downstream areas to compare the actual and predicted visual motion and update the internal model of the world.

5.7.3 Circuit mechanisms of depth representation in V1

Depth selectivity in V1 may be supported by long-range and local circuitry. Depth-selective responses may be generated in V1 *de novo* by integrating locomotion-related signals from the basal forebrain (Fu et al., 2014) and/or the secondary motor cortex (Leinweber et al., 2017). During locomotion, cholinergic innervations from the basal forebrain activate the nicotinic receptors of VIP neurons in V1 (Fu et al., 2014). The secondary motor cortex and the A42b region of the anterior cingulate cortex also provide a major input to L2/3 and L6 excitatory cells and PV interneurons in V1 (Leinweber et al., 2017). Alternatively, V1 neurons may inherit depth selectivity from upstream structures such as the thalamus or superior colliculus (SC). Subcortical regions that provide direct or indirect input to V1, such as dLGN, LP and superior colliculus (SC), have been shown to exhibit varying degrees of visuomotor integration. Locomotion onset can increase spontaneous and visually-driven responses of neurons in dLGN, albeit to a much less degree compared to V1 (Erisken et al., 2014). When mice were running in a VR environment with simple gratings and circles, the activity of a subset of dLGN boutons ($\sim 20\%$) in

V1 was found to be positively corrected with both running and visual motion speed, potentially facilitating the cooperative integration of these signals (Roth et al., 2016). On the contrary, many LP boutons in V1 displayed anti-correlated tuning to running and visual motion speed, potentially reflecting their role in signalling visuomotor mismatch (Roth et al., 2016). As for SC, minor locomotion-related modulation of neuronal activity was found in its superficial lamina (Savier et al., 2019). Given their ability to integrate visual and motor information, dLGN, SC and LP could serve as potential sites for depth computation from motion parallax before visual signals reach V1.

The multiplicative visuomotor integration observed in V1 pyramidal neurons can be supported by local excitatory-inhibitory circuitry. VIP neurons in the visual cortex, a group of interneurons that mainly inhibit SST inhibitory neurons (Jiang et al., 2015; Pfeffer et al., 2013), were found to increase their responses during running (Fu et al., 2014; Jackson et al., 2016; Reimer et al., 2014). In contrast, the activity of SST neurons are suppressed by locomotion (Fu et al., 2014). Therefore, it has been proposed that locomotion-related modulation of pyramidal neurons in V1 is achieved by VIP neurons conveying locomotion signals via disinhibition (Fu et al., 2014). A previous study in *Drosophila melanogaster* has suggested that disinhibition can support multiplicative integration of signals (Groschner et al., 2022). It was hypothesized that the detection of motion direction in T4 neurons in the fly optic lobe requires a multiplicative computation between delayed and non-delayed signals from adjacent photoreceptors, in order to effectively enhance responses to the preferred direction and suppress responses to the non-preferred ones (Haag et al., 2016; Strother et al., 2017). Groschner et al. (2022) further showed that this multiplicative computation is achieved by the coincidence of cholinergic excitation and the release from glutamatergic shunting inhibition (disinhibition) through conductance reduction. A similar mechanism may underlie the generation of depth-selective responses through multiplicative integration of optic flow speed and running speed signals. Excitatory inputs conveying optic flow speed information may coincide with a release from inhibition by activating VIP neurons carrying running information from VIP neurons, enabling multiplicative computation between running speed and optic flow speed tuning. This mechanism would require connections to be formed between specific excitatory and inhibitory neuronal pairs with opposite running speed tuning.

5.7.4 Closed loop coupling is required for the accurate representation of depth

When breaking the coupling between running and optic flow in the open loop condition, running speed preferences, optic flow speed preferences and the peak response amplitude did not change for individual neurons. This suggests that the conjunctive coding of running speed and optic flow speed does not depend on the closed loop coupling between running and optic flow feedback. However, it is much more difficult to decode

depth from the population activity in V1 in the open loop condition. Although the overall statistics of optic flow speed may provide a rough estimate of depth, accurate computation of depth requires access to information about the running speed that gives rise to the optic flow feedback. The importance of closed loop coupling for the accurate depth representation is reminiscent of previous research on the development of depth perception in cats, which demonstrates that closed loop coupling between active locomotion and the resulting visual feedback is essential for the development of normal depth perception (Hein et al., 1970). This also indicates that the computation of depth in V1 may rely more on running speed information provided by the efferent copy of motor command rather than on the reafference signal from sensory feedback.

5.7.5 Summary

In conclusion, I show that depth selectivity in V1 arises from conjunctive coding of running speed and optic flow speed. Within V1 population, neurons with similar depth preferences display a wide range of running speed preferences, enabling the representation of a wide range of depths when animals are running at any speed. Accurate representation of depth in V1 requires the closed loop coupling between locomotion and optic flow feedback. These results bring a new perspective on the role of locomotion-related modulation in visual processing, and suggest that the widespread locomotion-related modulation in V1 may serve as a foundation for inferring from motion parallax – a crucial computation for the daily lives of animals. In the next chapter, I will characterise the representation of depth across V1, and illustrate how V1 can create a depth map of the visual space from motion parallax.

Chapter 6

A depth map of the visual space from motion parallax in the mouse primary visual cortex

6.1 Authorship declaration

I, Yiran He, performed the majority of the experiments and analyses in this chapter. Antonio Colas Nieto performed $\sim 2/3$ of the two-photon recordings in 2 out of 7 GCaMP6f or GCaMP6s mice. Xavier Cano-Ferrer and George Constantinou (members of the Making Lab at the Francis Crick Institute) built parts of the customised electronics in the behavioural setup, including screen blanking (see Methods 2.4.8), lick detection, and the printed circuit board for synchronising camera and imaging triggers. This chapter built upon findings that I have previously disseminated in a preprint on bioRxiv (He et al., 2024).

6.2 Background

Building on the pioneering studies of Hubel and Wiesel (1959, 1962), visual responses of excitatory neurons in the mouse primary visual cortex (V1) have been extensively characterised, primarily using simplified, parametric, two-dimensional (2D) stimuli such as moving bars, drifting gratings and noise stimuli (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). These studies have led to several prevailing views regarding the properties of V1 neurons. It has been widely acknowledged that mouse V1 is organised retinotopically, with individual neurons responding to particular locations in the 2D visual field and neighbouring neurons representing adjacent areas in the visual field (Dräger, 1975; Kalatsky & Stryker, 2003; Wagor et al., 1980). Individual V1 neurons exhibit specific preferences for several key parameters of the visual scene, including orientation, spatial frequency, temporal frequency, contrast, and direction of motion (Andermann et al., 2011; Bonin et al., 2011; Dräger,

1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). These findings have given rise to a widely accepted view that V1 acts as a collection of Gabor filters on 2D retinal images (Bonin et al., 2011; Niell & Stryker, 2008).

Although 2D gratings are powerful tools to study the receptive fields (RFs) of V1 neurons, it imposes strict assumptions on the tuning properties of neurons. Recent studies have started to harness deep neural network to synthesize non-parametric stimuli based on natural images and have identified more complicated receptive fields in mouse V1 (Bashivan et al., 2019; Fu et al., 2024; Walker et al., 2019). Unlike monkey V1 neurons which are mostly excited by Gabor stimuli, RFs of mouse V1 neurons have a complicated organisation of subfields with distinct orientation preferences (Fu et al., 2024). However, regardless of parametric or non-parametric stimuli, previous research have rarely probed how V1 neurons respond to the third-dimension – depth – of the visual scene and whether depth is represented in RFs of V1 neurons. Using a virtual reality (VR) environment to present three-dimensional (3D) visual scenes, I have demonstrated in Chapter 4 that a majority of V1 neurons are selectively tuned to virtual depth of the stimuli. Given that V1 neurons display RFs at specific locations of the 2D visual field, I further hypothesize that depth-selective neurons respond to visual cues at specific retinotopic locations and at specific virtual depths. In other words, the activity of depth-selective neurons in V1 can be described by 3D RFs.

If V1 neurons have 3D RFs, another important question to consider is how these RFs are spatially distributed across V1. While RFs of V1 neurons are expected follow retinotopic gradients along the dimensions of azimuth and elevation, their distribution along the dimension of depth remains unknown. Multiple studies have shown that neuronal populations in V1 and the surrounding higher visual areas (HVAs) have distinct preferences for visual motion speed mapped by grating stimuli (Andermann et al., 2011; Garrett et al., 2014; Glickfeld & Olsen, 2017; Marshel et al., 2011). While V1 neurons show diverse tuning to visual motion speed covering a wide range, HVAs exhibit more specific preferences, with anterior HVAs (e.g. AM, A, RL and AL) preferring fast-moving stimuli and posterior HVAs (e.g. PM) preferring slow-moving stimuli. Since optic flow speed preferences are indicative of depth preferences (Figure 5.4A), I hypothesize that V1 consists of heterogenous populations preferring distinct depths, with neurons preferring diverse depths intermingled with each other. Meanwhile, different HVAs may display specialisation in representing near or far depths.

6.3 V1 neurons have three-dimensional receptive fields

To test the hypothesis that neurons are tuned to specific retinotopic locations and specific virtual depths, I first reconstructed the stimuli presented during each imaging frame (Figure 6.1A). Then to characterise the RFs of individual neurons, I fitted neuronal responses ($\Delta F/F_0$) as a function of the reconstructed stimuli in each imaging frame (S)

and the RF of each neuron (**b**) using a linear model, such as $\Delta F/F_0 = S\mathbf{b}$. This model accounted for the modulation of neuronal responses by retinotopic locations in azimuth and elevation as well as virtual depth, while maintaining smoothness of the receptive field along the azimuth and elevation dimensions and the tuning to virtual depth (see Methods 2.4.11). Only the stimuli in the right visual field of the mice was taken into account (Figure 6.1A), as imaging was conducted in the left visual cortex.

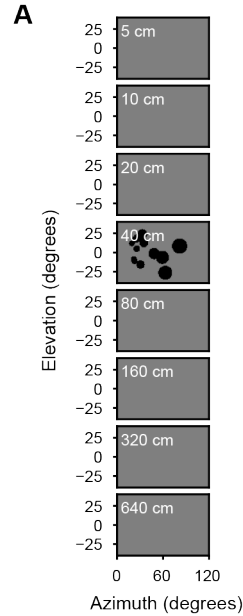


Figure 6.1: **Schematics of linear receptive field model.** (A) An example reconstructed stimulus frame. Stimuli were downsampled to 5 degrees for analysis.

I found that many neurons responded to visual stimuli at specific retinotopic locations, and their spatially-selective activity was modulated by virtual depth (Figure 6.2A-D). To determine the proportion of neurons that have identifiable receptive fields, I estimated the null distribution for RF coefficients for each neuron by fitting the RF using the stimuli in the left visual field (RF_{left}). Then I compared it with the RF estimated using the right visual field (RF_{right}), and determined the neurons with significant RFs as those with a maximum value of the RF_{right} which was 6 standard deviations above the mean of RF_{left} . Across recording sessions, the majority (66.3%, 20,554 / 31,013) of depth-selective neurons displayed identifiable receptive fields (Figure 6.2E). In other words, V1 neuronal responses can be characterised by 3D RFs during active navigation through a 3D VR environment.

I then explored the spatial organisation of the 3D RFs within recording sessions. As expected, the spatial location of RFs recorded within the same field of view was clustered in a particular region of the visual field (Figure 6.3A-B). The preferred azimuth and elevations of neurons followed retinotopic gradients (Figure 6.3A-B,D-E), validating the methods of RF estimation. However, the distribution of virtual depth preferences exhibited a salt-and-pepper pattern, with neighbouring neurons often preferring disparate virtual depths (Figure 6.3C,F).

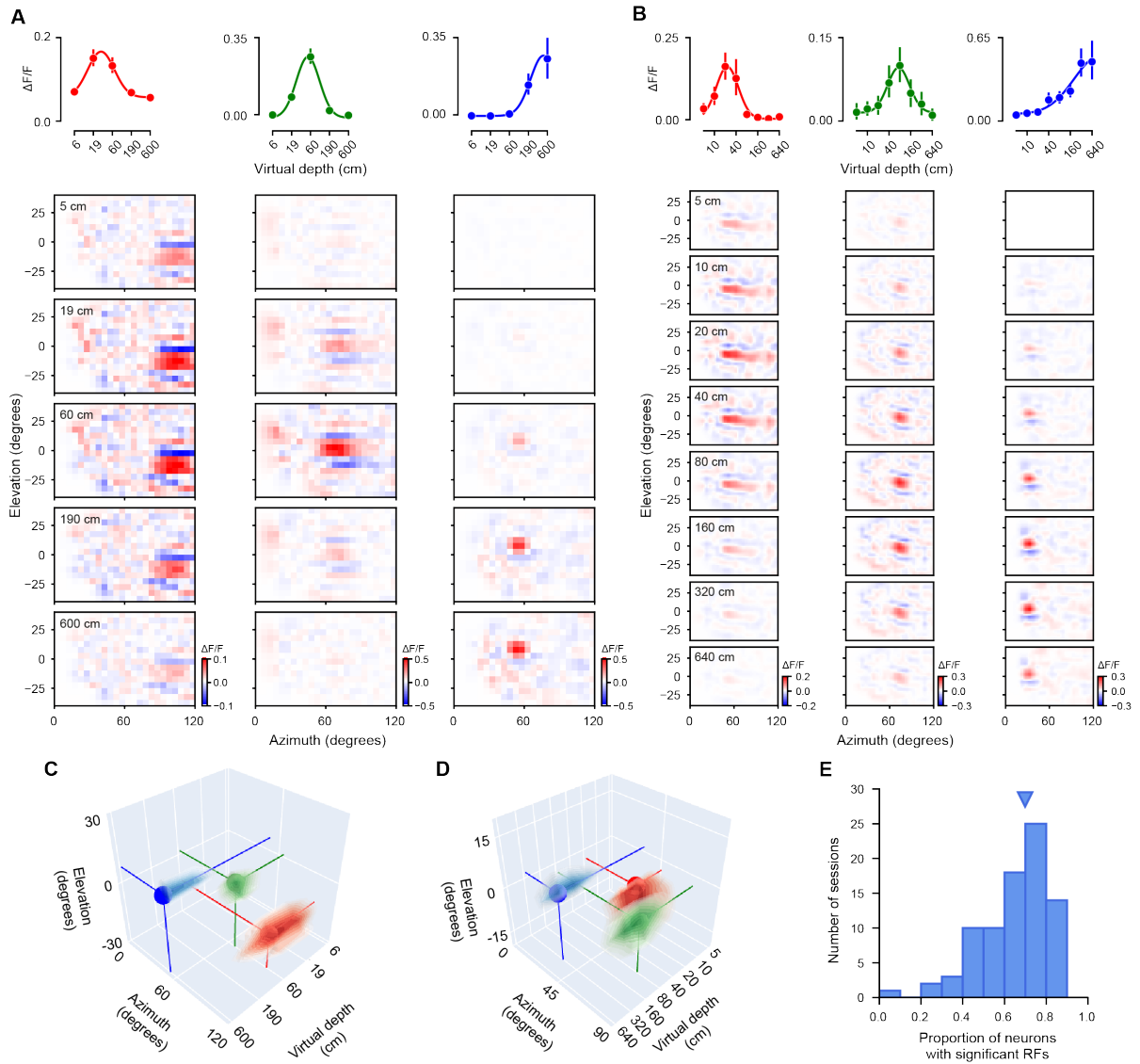


Figure 6.2: **Depth-selective neurons have three-dimensional receptive fields.** (A-B) Receptive fields of 6 example V1 neurons in 2 sessions as a function of virtual depth. (C-D) Receptive fields of the example V1 neurons in A-B visualized in 3D. The rendered volume represents the receptive field locations with responses greater than half of maximum for each cell. (E) Proportion of neurons with significant receptive fields across imaging sessions (N = 85 sessions). Triangle indicates the median value.

6.4 The representation of depth across V1

I then characterised the spatial distribution of retinotopic and depth preferences of neurons across V1. Neurons within V1 were identified by mapping the boundaries between V1 and HVAs using widefield calcium imaging (Figure 6.4A, see Methods 2.4.9). To align the cortical locations of neurons across all animals, I first identified the spatial location of recorded neurons within V1 in individual animals, by determining the location of neurons within each field of view based on cell masks and the location of each field of view within the window based on blood vessel patterns. Next, I assumed that RF azimuths and elevations followed the same linear gradient in V1 across animals, and found the translational offsets between windows of different animals. The spatial locations of

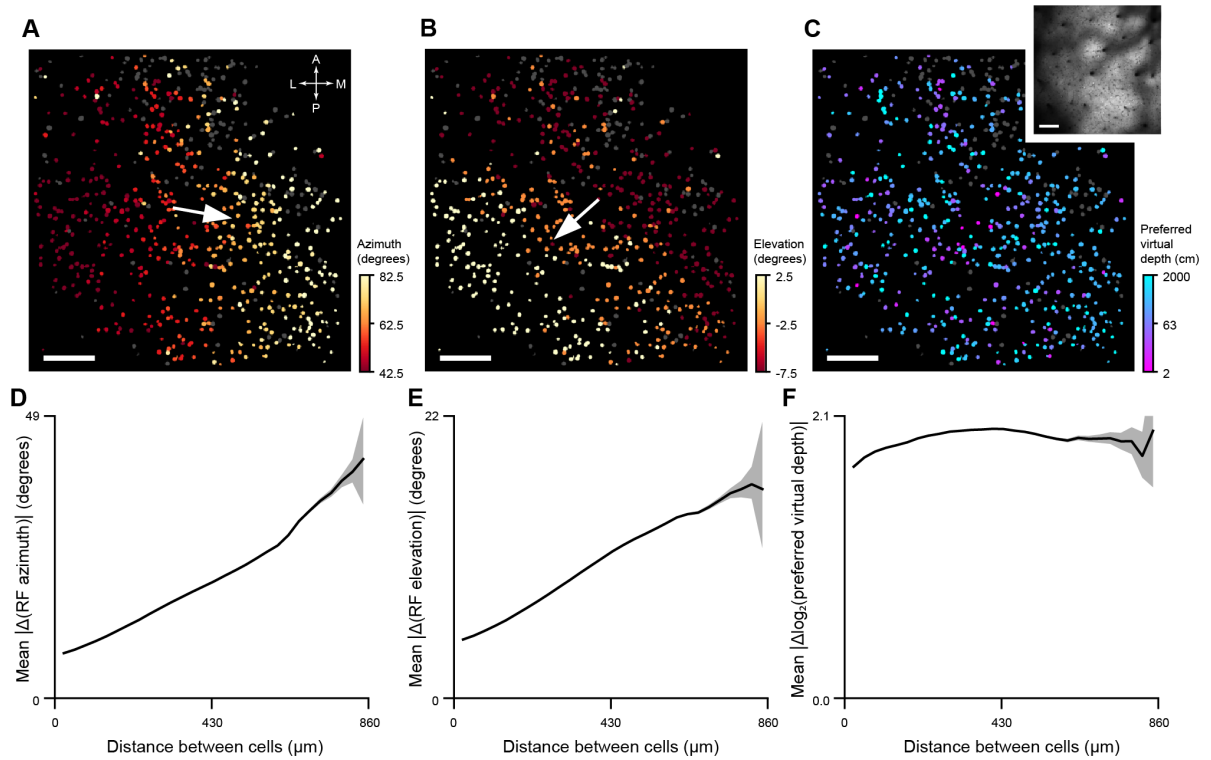


Figure 6.3: Depth preferences display a salt-and-pepper distribution. (A-C) Preferred azimuth (A), elevation (B), and depth (C) of depth-selective neurons across the field of view in an example imaging session. A – anterior; P – posterior; M – medial; L – lateral. Scale bar – 100 μm . Arrow – direction of gradient of increase in preferred azimuth or elevation. Inset – mean fluorescence image. (D-F) Mean pairwise distance of RF azimuth (D), RF elevation (E) or log-preferred depth (F) between depth-tuned neurons with significant RFs binned by the distance ($N = 20,554$ neurons from 85 sessions). Pairs of neurons within 10 μm from each other were excluded to avoid the effects of fluorescence contamination. Shading – 95% confidence interval computed by bootstrap resampling neuron pairs.

neurons within V1 across animals were aligned to a reference animal (see Methods 2.4.11).

After aligning the cortical locations of neurons across all animals, I visualised the distribution of preferred azimuth, elevation and virtual depth across V1. As expected, preferred azimuth and elevation of neurons across V1 still followed retinotopic gradients (Figure 6.4B-C). Preferred virtual depths of neighbouring neurons were intermingled in general, yet displaying certain biases in spatial distribution (Figure 6.4D). Notably, posterior V1 was enriched with neurons preferring near virtual depths (Figure 6.4D), and the virtual depth preferences were negatively correlated with the anterior-posterior locations of depth-selective neurons in V1 ($r = -0.239$, $p = 0.0007$), but were not significantly correlated with their lateral-medial locations ($r = 0.094$, $p = 0.079$).

I further analysed the relationship between virtual depth preferences and the RF location of individual neurons. In previous analysis, virtual depth was defined as the virtual distance to the sphere stimuli when the spheres passed 90 degrees of azimuth from the animal's perspective. However, the depth of the stimuli varies at different retinotopic locations due to the cylindrical design of the stimuli, and the spheres that were not

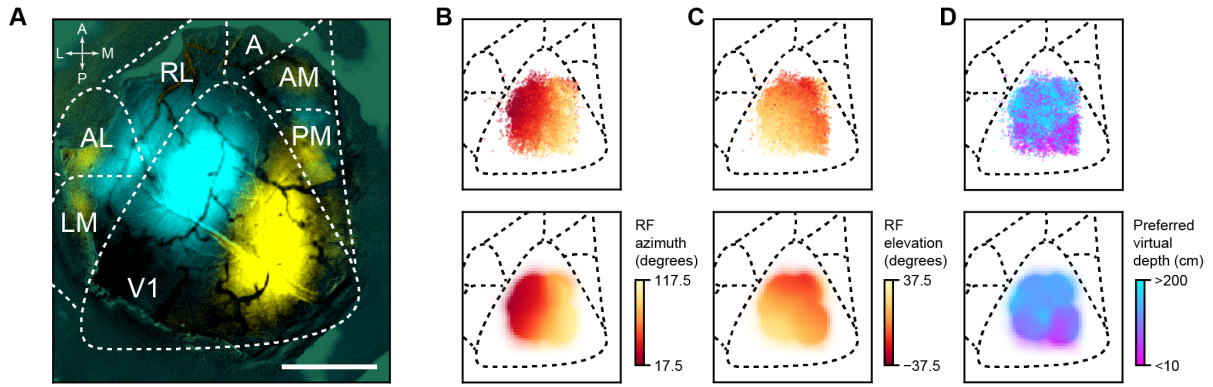


Figure 6.4: The representation of depth across V1. (A) Location of V1 and higher visual areas identified using widefield calcium imaging. Cyan and yellow indicate responses to grating patches centered at 30 and 90 degrees of azimuth. Scale bar – 1 mm. LM, lateromedial; AL, anterolateral; RL, rostrolateral; A, anterior; AM, anteromedial; PM, posteromedial. (B-D) Top – spatial distributions of receptive field (RF) azimuths (B) and elevations (C), and preferred virtual depths (D) of depth at the location of individual neurons; bottom – RF azimuth, elevation and preferred virtual depth smoothed with a Gaussian kernel (N = 20,554 neurons from 85 sessions).

at 90 degrees azimuth were further away than the virtual depth value of the trial. To correct for this effect, I adjusted the preferred depth of individual neurons by calculating the distance of spheres to the animal at the neuron's preferred retinotopic location (see Methods Eq.2.4.29, Figure 6.5C). This correction had the largest effect at small azimuths, where spheres were ~ 2 times further compared to the spheres at the closest point (at 90 degrees azimuth). However, this difference was substantially smaller than the >100 fold range of depths sampled across trials.

By visualising the distribution of virtual depth preferences as a function of RF location, I found that neurons tuned to near depths were over-represented in the upper and lateral visual field (Figure 6.5A-B,D-E), and this effect was significant regardless of the correction for preferred depth described above (corrected, $p = 6.51 \times 10^{-5}$; uncorrected, $p = 0.00761$). Despite the biased distribution of neurons preferring near depths, neurons tuned to a wide range of depth values could be observed in all retinotopic locations (Figure 6.5B,E). Consistent with what was observed in individual fields of view, virtual depth preferences were distributed in a salt-and-pepper manner throughout the visual field sampled by the stimuli (Figure 6.5B,E). When separating depth-selective neurons into 3 populations - near- (<20 cm), mid- (20-100 cm), and far (>100 cm) preferring cells, the 3 populations of neurons were intermingled (Figure 6.5B,E). Overall, these results suggest that V1 contains a depth map of the visual space from motion parallax, with neuronal populations preferring distinct depths intermingled with each other. This organisation of depth preferences allows neurons responding to similar locations in the visual field to encode a broad range of distances from various objects in the visual scene.

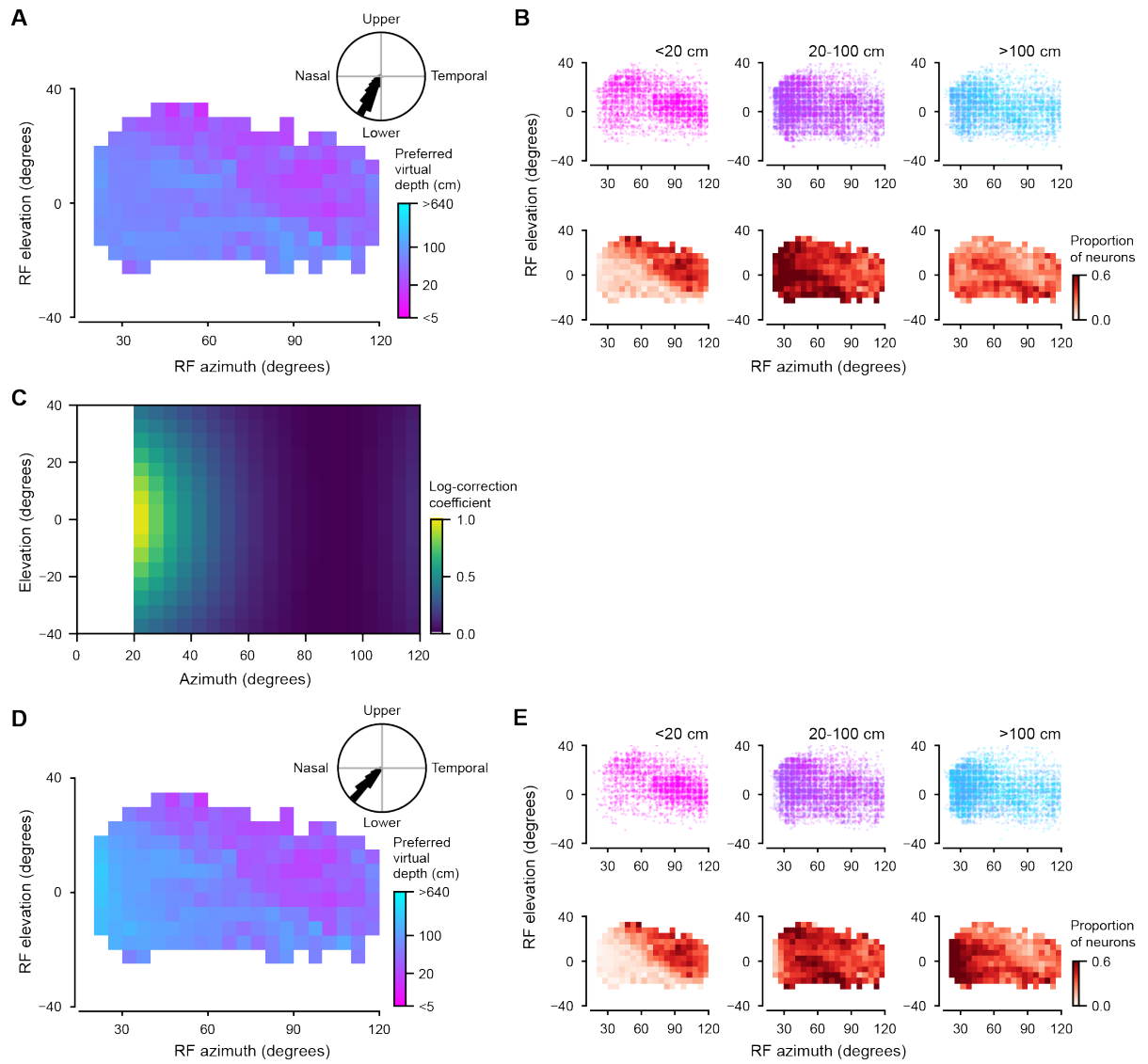


Figure 6.5: Relationship between depth preferences and RF locations. (A) Median uncorrected preferred virtual depth as a function of RF location ($N = 20,554$ neurons from 85 sessions). Inset – direction of the gradient of preferred virtual depth across the visual field estimated across hierarchical bootstrap samples of the dataset. RF locations were estimated with 5 degree resolution and was plotted with 2 degrees of jitter. (B) RF locations of depth-selective neurons tuned to near (<20 cm, $N = 4,518$ neurons), intermediate (20-100 cm, $N = 10,325$ neurons) and far (>100 cm, $N = 5,711$ neurons) virtual depths without correcting for RF location. Bottom – proportion of depth-selective neurons in each category as a function of RF location. (C) Log-transformed correction coefficients as a function of RF location. (D) Median corrected preferred virtual depth as a function of RF location ($N = 20,554$ neurons from 85 sessions). Only locations with at least 20 neurons are shown. Inset – direction of the gradient of preferred virtual depth across the visual field estimated across hierarchical bootstrap samples of the dataset. (E) RF locations of depth-selective neurons tuned to near (<20 cm, $N = 3,753$ neurons), intermediate (20-100 cm, $N = 8,412$ neurons) and far (>100 cm, $N = 8,389$ neurons) virtual depths (top) after correcting for RF location, and proportion of depth-selective neurons in each category as a function of RF location (bottom).

6.5 Discussion

6.5.1 A depth map of the visual space in V1

Neuronal responses in V1 have been traditionally studied with simplified, parametric, 2D stimuli such as moving bars, drifting gratings and noise stimuli (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). This has led to a popular view that V1 neurons can be modelled as a collection of Gabor filters with specific preferences to spatiotemporal frequencies and orientations of visual stimuli (Andermann et al., 2011; Bonin et al., 2011; Gao et al., 2010; Hubel & Wiesel, 1962; LeDue et al., 2012; Niell & Stryker, 2008; Skottun et al., 1991). Here I recorded V1 neuronal responses when mice were navigating in a 3D VR environment. Then, I analysed their RFs by reconstructing the visual stimuli at each imaging frame, and fitted the neuronal responses with a linear model incorporating retinotopic locations in azimuth and elevation as well as virtual depth. While a large proportion of V1 neurons show selectivity towards virtual depth from motion parallax as demonstrated in Chapter 4, the majority of them ($\sim 66\%$) displayed clear canonical 2D RFs as characterised in previous studies (Dräger, 1975; Kalatsky & Stryker, 2003; Wagor et al., 1980). They responded to particular locations in the 2D visual field as well as specific virtual depths. In other words, V1 neuronal responses can be characterised by 3D RFs during active locomotion. The proportion of depth-selective neurons with detectable RFs represents a lower bound, as neurons with responses at lower signal-to-noise ratios may have RFs that are challenging to detect. Overall, the discovery of 3D RFs of V1 neurons supports a novel view of how mouse V1 represents the visual scene in actively moving animals. These results demonstrate that motion parallax generates a depth map of visual space in V1, with distinct neuronal populations responding to near versus far visual cues.

The preferred retinotopic locations of the neurons followed expected retinotopic gradients (Dräger, 1975; Kalatsky & Stryker, 2003; Wagor et al., 1980), validating the method of estimating RFs. Moreover, the RFs of many neurons in azimuth and elevation displayed Gabor-like features, showing separated yet aligned ON and OFF subregions (Figure 5.1A-B) as previously characterised in V1 simple cells (Bonin et al., 2011; Hubel & Wiesel, 1962; Niell & Stryker, 2008; Skottun et al., 1991). Nevertheless, it is worth noting that RF estimation is limited by the statistics of the visual stimuli. Albeit being in three dimensions, my VR stimuli only consisted of black spheres and gray background moving in the direction of the animal's locomotion, lacking the complex statistics that would appear in natural scenes. As a result, I did not observe the complicated RF subregions as mapped by Fu et al. (2024) using natural images. In future experiments, the RFs of V1 neurons could be mapped using more naturalistic 3D environments that incorporate objects with diverse shapes, contrasts, and colours, reflective of the visual scenes found in nature.

6.5.2 The distribution of depth representation across V1

In contrast of the gradual retinotopic gradients, the distribution of virtual depth preferences within V1 exhibited a salt-and-pepper pattern. Neurons responding to similar retinotopic locations displayed a wide range of virtual depth preferences, which potentially allows a broad range of depth values to be represented in every part of the visual field. This result is consistent with previous studies which show that V1 neurons have a diverse range of visual motion speed preferences, whereas HVAs have more specialised speed preferences when mapped with 2D (Andermann et al., 2011; Garrett et al., 2014; Glickfeld & Olsen, 2017; Marshel et al., 2011). As preferred depth is inversely correlated with preferred optic flow speed, I hypothesize that anterior HVAs (e.g. AM, A, RL and AL) that prefer fast-moving stimuli (Andermann et al., 2011; Marshel et al., 2011) will specialise in representing near depths. On the contrary, posterior HVAs (e.g. PM) prefer slow-moving grating stimuli (Andermann et al., 2011; Marshel et al., 2011) and are proposed to provide information about visual landmarks to the navigation system for the formation of spatial representation of the environment (Roth et al., 2012). Therefore, I hypothesize that posterior HVAs will bias towards representing far depths.

Although nearby V1 neurons show disparate virtual depth preferences, the distribution of depth preferences in V1 is not homogenous. The most notable trend is the over-representation of near depths in the upper lateral visual field. This salient trend cannot be solely attributed to the bias inherent to the layout of the virtual environment. The virtual depths of objects presented at different locations of the visual field were symmetrical along the horizontal axis (0 degree elevation) and along the vertical axis of 90 degree azimuth, yet the distribution of virtual depth preferences were asymmetrical across these two axes.

The distribution of depth representation from motion parallax is different compared to depth preferences characterised by binocular disparity in binocular V1. Nearby depths are over-represented in the upper visual field of V1. In contrast, the lower visual field of V1 binocular zone (and RL binocular zone which primarily covers the lower visual field) tend to represent nearer depths (La Chioma et al., 2019). This distinction may indicate the ethological relevance of monocular and binocular zones in V1. Binocular zone was found to specifically support tracking preys during hunting behaviour, whereas mice with only monocular vision performed poorly in hunting (Johnson et al., 2021). Mice seldom use motion parallax as a depth cue during prey pursuit, as they tend to coordinate their head and eye movements to keep their prey in a retinal region with minimal optic flow (Holmgren et al., 2021). On the contrary, the monocular zone of V1 may specialise in detecting predators in the periphery, when high sensitivity to visual motion is critical and binocular information may not be available. The distribution of depth preferences in V1 may be inherited from the bias in the distribution of retinal ganglion cells (RGCs) in the retina. RGCs are not evenly distributed across the mouse retina, with the highest density of RGCs in the ventronasal retina (Dräger & Olsen, 1981) corresponding to the upper lateral visual field. W3 cells, a type of RGCs with the smallest RF that are

densely enriched in the ventral retina, are specialised for detecting objects moving against a featureless or stationary background (Zhang et al., 2012). The ability of detecting local motion by W3 cells is considered as an alarm mechanism for detecting predators (Zhang et al., 2012). However, it was also proposed that the small receptive field of W3 cells may be designed for the detection of far predators covering a smaller angular subtense (Zhang et al., 2012). Further investigation is needed to determine the ethological relevance of the distribution patterns of depth representation in V1.

6.6 Summary

In conclusion, I have presented a revolutionary view of how V1 is functionally organised to process visual scenes. Instead of the traditional view that V1 acts as a set of Gabor filters on 2D retinal images, I show that motion parallax generates a depth map of 3D visual space in V1, with distinct population of neurons representing near or far visual cues. Nearby neurons prefer disparate virtual depths, which enables a broad range of depths being represented in each part of the visual field. Interestingly, computer vision algorithms use a similar approach, generating an estimation of depth for each pixel of the visual scene (Du et al., 2020; Fu et al., 2019; Gimeno et al., 2013; Liao et al., 2020; You et al., 2020). The strategy of encoding a depth map of the visual scene allows both biological agents and artificial agents to navigate in complex 3D environments.

Chapter 7

Discussion

7.1 Summary

Understanding the structure of the three-dimensional (3D) environment is crucial for animals' daily life. Many animals have an innate ability to perceive depth (Walk & Gibson, 1961). Nevertheless, how 3D scenes are parsed by the visual system for depth estimation remains poorly understood. Motion parallax is an important monocular cue for depth perception that relies on the visual motion of objects as a result of the observer moving in the environment. As the observer move through the environment, depth can be estimated by comparing the speed of locomotion and the speed of resulting visual motion. Motion parallax alone is sufficient for humans to generate a sense of depth (Rogers & Graham, 1979), and it is especially valuable for animals without extensive binocular overlap (e.g. rodents). How neurons in the visual system support depth perception from motion parallax remains to be systematically characterised.

My project focuses on the depth perception of mice and aimed to examine two main questions:

1. Is depth perception innate for mice?
2. How does the primary visual cortex (V1) support depth perception from motion parallax in mice?

In Chapter 3, I presented a customized version of the visual cliff test (Walk & Gibson, 1961) as a behavioural assay to systematically characterise depth perception in mice. I first demonstrated that mice can discriminate depth reliably on the visual cliff test using visual cues, and specifically non-pictorial visual cues. Then, I showed that depth perception does not require any prior visual experience in mice, by comparing the performance of mice born and reared in the dark for 3 months to their counterparts raised in normal lighting. Finally, I presented the preliminary results showing that superior colliculus (SC) – one of the main subcortical structures providing inputs to the visual cortex – is not required for depth perception, by quantifying the performance of mice on the visual cliff under optogenetic inhibition of the superior colliculus. The innate nature of depth perception

for mice, along with many other animals, highlights the functional significance of depth perception throughout evolution. The neural circuits supporting depth perception are thus likely conserved through evolution across different species, and are likely to involve subcortical structures in addition to the visual cortex, which was proven indispensable for rats (Meyer et al., 1966; Meyer, 1963). However, SC does not seem to be required by mice to perform depth discrimination tasks. Therefore, other subcortical inputs such as thalamic nuclei (e.g. dLGN, LP) as well as the visual cortex may be the essential neural substrates for depth perception in mice.

Next, in Chapter 4 to 6, I explored the neural mechanism underlying depth perception from motion parallax, focusing on V1. V1 neurons are widely modulated by locomotion (Keller et al., 2012; Niell & Stryker, 2010; Saleem et al., 2013), and can integrate both visual motion and locomotion-related signals (Keller et al., 2012; Saleem et al., 2013). Therefore, I hypothesized that V1 neurons can integrate both visual motion and locomotion-related signals to compute motion parallax for depth estimation. To test this hypothesis, I designed a virtual reality (VR) environment where spheres were presented at various virtual distances to the mice. I then recorded neuronal responses in L2/3 of V1 using two-photon microscopy while head-fixed mice navigating in the VR environment. The parameters of the VR environment were precisely controlled so that motion parallax was the only cue available for depth estimation. I found that over half of L2/3 neurons in V1 were selectively tuned to virtual depth. This depth selectivity was driven by motion parallax, with little neuronal responses during stationary periods of visual stimulus presentation. Depth selectivity of neurons were also unaffected by the size of the stimuli despite expected changes in response magnitude to different stimulus sizes (Adesnik et al., 2012; Dipoppa et al., 2018).

I then investigated how depth selectivity arises from locomotion and visual motion signals in Chapter 5. I compared different computational models of depth-selective responses being generated by running speed tuning alone, optic flow speed tuning alone, or the integration of both running speed and optic flow speed tuning. I show that depth selectivity arises from the specific conjunction of running speed and optic flow speed tuning. The preferred virtual depth is determined by the ratio of preferred running speed and preferred optic flow speed. When breaking the coupling between running and optic flow feedback, running speed preferences, optic flow speed preferences or the peak depth-selective responses of neurons do not change, indicating that the conjunctive coding of individual neurons does not rely on closed loop visuomotor coupling. However, on a population level, closed loop coupling enables the accurate representation of depth encoded by population activity of V1 neurons. In addition, I compared the speed preferences of V1 neurons mapped by the closed loop VR stimuli during active locomotion of the mice and two-dimensional (2D) drifting gratings during passive viewing. Traditionally, visual responses of V1 neurons have been studied using simple parametric 2D stimuli, such as bars or drifting gratings, while animals passively observe the stimuli (Andermann et al.,

2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). Despite the small sample size, I found a positive correlation between the speed preferences of neurons mapped using 3D VR stimuli and 2D drifting gratings. This suggests that the speed preferences of V1 neurons remain consistent across conditions of passive visual stimulation and self-generated visual flow.

Finally in Chapter 6, I characterised the receptive field (RF) of V1 neurons and the distribution of depth representation in V1. RFs of V1 neurons have been traditionally characterised by 2D stimuli such as gratings or natural images, leading to a view that V1 is a collection of spatiotemporal filters where each neuron respond to a specific location in the 2D visual field and stimuli with specific spatiotemporal frequencies, orientations and moving directions (Andermann et al., 2011; Bonin et al., 2011; Gao et al., 2010; Hubel & Wiesel, 1962; LeDue et al., 2012; Niell & Stryker, 2008; Skottun et al., 1991). By modelling RFs of depth-selective neurons in V1, I found that these neurons responded to specific retinotopic locations in azimuth and elevation as well as specific virtual depths, displaying 3D RFs during active locomotion. While azimuth and elevation preferences exhibited expected retinotopic gradients in V1, virtual depth preferences displayed a salt-and-pepper pattern, with nearby neurons preferring disparate virtual depths. This enables V1 to represent a wide range of depths at each retinotopic location. Across V1, neurons preferring different virtual depths were generally intermingled, with some biases in the upper lateral visual field overrepresenting near depths. The distribution of depth preferences may reflect the ethological relevance of different parts of the visual field in natural behaviours such as hunting. Overall, motion parallax creates a depth map of visual space in V1, with distinct population of neurons representing near or far depths.

7.2 Depth perception in naturalistic environments

Building on the pioneering work from Hubel and Wiesel (1959, 1962), existing studies have predominately examined V1 neuronal activity using passive visual stimulation consisting of simplified parametric 2D stimuli, characterising the selectivity of V1 neurons for specific visual features such as spatiotemporal frequency and orientation (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). However, such stimulation differ significantly from the visual experience of animals in the real world. Natural environments usually contain diverse and complex three-dimensional structures. Moreover, animals predominantly use vision as an active sense – they acquire visual information during active exploration of the environment to guide essential behaviours such as navigation and foraging (Skyberg & Niell, 2024). Recent studies have started to harness deep neural network to synthesize non-parametric stimuli with statistics are more akin to natural images, and have identified receptive fields (RFs) of V1 neurons that are remarkably different from canonical Gabor-like RFs in mice (Bashivan et al., 2019; Fu et al., 2024; Walker et al., 2019). In this project, I focused on

incorporating the third-dimension – depth – into the visual environment, and developed a VR system that allows animals to experience self-generated visual flow during active locomotion while maintaining precise control over stimulus parameters. However, this VR environment was intentionally simplified so that motion parallax depth cues could be studied in isolation. Real-world environments typically contain much richer optic flow patterns and additional cues that can contribute to depth perception.

7.2.1 Object motion

All objects were stationary in the VR environment I designed. However in the real world, objects can be either stationary or moving. Therefore, visual motion can be generated by both self-motion of the observer and object motion. How would depth-selective neurons in V1 respond to moving objects and would V1 population still be able to encode depth accurately with the presence of object motion?

For primates and humans, object motion can be determined by the conflicts between binocular disparity and motion parallax cues. When primates or humans fixate their eyes to focus on a specific point in space and translate in the environment, objects closer than the fixation plane appear to move in the opposite of self-motion, whereas objects farther than the fixation plane appear to move at the same direction of self-motion (Nadler et al., 2008). Therefore, object motion in the opposite direction of the direction of visual flow predicted from its depth sign (near or far) can create conflicts between motion parallax and disparity cues. For example, if an object farther than the fixation plane moves towards the opposite direction of self-motion at a faster speed than self-motion, motion parallax cues will indicate that the object is closer than the fixation plane. The neural mechanisms of computing object motion from the conflicts between multiple depth cues were proposed by Nadler et al. (2013), where they measured the neuronal activity in macaque area MT in response to either motion parallax or binocular disparity cues. Many neurons (60%) showed matched preferences of depth sign (near or far) in response to each cue (congruent cells). Surprisingly, 40% of MT neurons preferred opposite depth signs when depth was specified by binocular disparity vs, motion parallax cues (opposite cells). While congruent cells may be responsible for integration of depth cues for more accurate depth perception, opposite cells may signal the conflict between motion parallax and disparity cues (Nadler et al., 2013). When objects moving in the opposite direction of its depth sign, opposite cells may be more responsive compared to congruent cells, and thus support the detection of object motion during self-motion (Kim et al., 2016; Nadler et al., 2013). This hypothesis is supported by a subsequent study which showed that opposite cells were more responsive to moving objects than stationary objects compared to congruent cells (Kim et al., 2022).

For mice and many other animals without an extensive binocular overlap (Heesy, 2004), the majority of the visual information comes from the monocular visual field. Object motion speed can be calculated from the population activity of depth-selective neurons in V1 via conjunctive coding of running and optic flow speed. Assuming that

optic flow speed can be derived from the ratio between self-motion speed (S) and depth (d), and that total optic flow speed (V) is a result of combining optic flow generated by self-motion speed (S) and object motion speed (O), we can express the total optic flow speed V as:

$$V = \frac{S}{d} + \frac{O}{d}. \quad (7.2.1)$$

Assuming that object motion and depth are constant for a short period of time, V can be expressed as a linear function of the self-motion speed plus a constant (b):

$$V = \frac{1}{d}S + \frac{O}{d} = \frac{1}{d}S + b, \quad (7.2.2)$$

where the slope of the linear equation is the inverse of depth ($\frac{1}{d}$) and the intercept is the ratio between object motion speed and depth ($\frac{O}{d}$). Therefore, object motion can be computed by downstream areas through population activity of V1 depth-selective neurons as animals sample the environment at various locomotion speeds.

7.2.2 Free navigation

In my experiments, mice were head-fixed and were running in a single direction on the treadmill. Head-fixed mice display much less eye movements compared to freely moving mice (Meyer et al., 2020), and lack vestibular feedback for estimating their self-motion. For freely moving mice performing naturalistic navigation in the environment, mice exhibit rich head and eye movement patterns – they make non-conjugate eye movements compensatory to head tilts (pitch and roll) for visual field stabilisation, as well as conjugate eye movements in a "saccade and fixate" manner in response to horizontal head rotation (yaw). Moreover, they are able to move in various directions so that optic flow from many directions can be experienced. Therefore, it is important to investigate how depth is encoded in V1 during free navigation, when the mice experience rich head and eye movement patterns as well as optic flow in various directions, and whether vestibular inputs contribute to depth selectivity in V1. This question could be addressed through recording V1 activity using electrophysiology while the mice are freely navigating in an arena projected with an VR environment with objects at various distances. To accurately estimate the retinal image in freely-moving mice, momentary head and eye movements would need to be tracked during the recording. Thanks to the newly developed recording technology with miniature head-mounted cameras (Meyer et al., 2018; Michaiel et al., 2020; Parker, Abe, Leonard, et al., 2022; Sattler & Wehr, 2021), visual scene from the perspective of an freely-moving animal could be reconstructed by combining the recorded eye movements and head movements.

7.2.3 Naturalistic scenes

Our VR environment isolated motion parallax as the only cue for depth by eliminating any cues related to binocular disparity, textures, shading, dynamic perspective (spheres do not change when viewed from different angles), etc. In the nature, there are rich binocular and monocular cues generated by objects with complex shapes, colours and texture patterns with various spatial frequencies. Moreover, recent studies have identified more complicated receptive fields in mouse V1 probed by natural images compared to the canonically-identified Gabor-like receptive fields (Bashivan et al., 2019; Fu et al., 2024; Walker et al., 2019). To explore how depth can be represented in V1 and how different depth cues can be integrated during real-world navigation, it is essential to design more naturalistic environments that resemble natural scenes, incorporating objects with diverse colours, shapes and texture patterns.

In summary, further experiments could focus on characterising the representation of depth in V1 using environments that closely mimic real-world scenes. This approach could involve allowing mice to navigate freely, incorporating naturalistic settings with richer depth cues – such as binocular and pictorial cues, as well as objects at varying depths – and introducing elements such as object motion and passive motion.

7.3 A depth map in the visual cortex

7.3.1 Depth representation across the visual cortex

Chapter 5 - 6 have characterised the distribution of depth-selective neurons in L2/3 of V1. V1 L2/3 populations have a wide range of depth preferences, with neurons preferring similar depths tuned to a variety of combinations of optic flow speed and running speed. The mixture of optic flow and running speed tuning can serve as basis functions for V1 L2/3 populations to encode the depth of objects when animals are running at various speeds. In downstream areas such as L5/6 of V1 or HVAs, neurons might pool the inputs from V1 L2/3 neurons with similar depth preferences to form tuning to depth that is invariant to running or optic flow speed.

Furthermore, depth preferences are largely intermingled in V1 – neighbouring neurons in V1 are tuned to disparate depths. Although a bias towards near depth representation can be observed in the area corresponding to the upper lateral visual field, neurons tuned to near, intermediate or far depths can be found across V1. This allows a wide range of depth values to be encoded at any retinotopic location, with the activity of certain regions emphasising salient stimuli. This result is also consistent with previous studies which mapped the vision motion speed preferences across the visual cortex and showed that V1 neurons have a diverse range of visual motion speed preferences, whereas HVAs have more specialised speed preferences (Andermann et al., 2011; Garrett et al., 2014; Glickfeld & Olsen, 2017; Marshel et al., 2011; Roth et al., 2012). Anterior HVAs including

AM, A, RL and AL were found to prefer fast-moving grating stimuli, whereas posterior HVAs such as PM were found to be more responsive to slow-moving stimuli. Due to the low sensitivity of motion, PM was also proposed to provide information on visual landmarks to the navigation system such as RSC to support spatial perception (Roth et al., 2012). Therefore, HVAs may also display functional specialisation in depth encoding, with anterior areas representing near depths and posterior areas representing far depths. This functional specialisation may be supported by specific projections from V1. For example, V1 neurons projecting to AL and PM are largely non-overlapping (Han et al., 2018; Kim et al., 2018). The tuning of V1 neurons also matches with the tuning of their target regions – AL projection neurons prefer high speed stimuli whereas PM projection neurons prefer high spatial-frequency which corresponds to lower speed (Kim et al., 2018). Therefore, V1 populations that prefer near depths may project to near-preferring HVAs such as AL, whereas V1 populations preferring far depths may project to far-preferring HVAs such as PM, generating specialisation in depth representation and visual functions in HVAs.

Further characterisation of depth preferences across the visual cortex can be achieved by recording depth-selective responses across different HVAs and different cortical layers. To examine the circuit mechanisms underlying the depth map across the visual cortex, *in vivo* imaging can be combined with retrograde labelling, using either virus injections at HVAs (Kim et al., 2018) or high-throughput *in situ* sequencing methods (Zhang et al., 2024). This approach can help determine whether potential specialisations in depth preferences within HVAs arise from specific V1 projections.

7.3.2 A novel perspective of V1

The neuronal activity of V1 has been traditionally probed using parametric and 2D stimuli such as moving bars, gratings and noise stimuli (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008). These stimuli allow us to systematically probe the neuronal responses to various key parameters of visual scenes such as spatial frequency, temporal frequency, luminance, orientation and retinotopic locations. However, these studies have resulted in a widely held view that V1 acts as a collection of Gabor filters, enabling the detection of specific spatiotemporal frequencies and orientations at specific locations in the 2D visual field (Andermann et al., 2011; Bonin et al., 2011; Gao et al., 2010; Hubel & Wiesel, 1962; LeDue et al., 2012; Niell & Stryker, 2008; Skottun et al., 1991). By introducing the third dimension – depth – into the visual stimuli while mice were navigating in a VR environment, I demonstrated that V1 neurons not only respond to specific locations in azimuth and elevation, but also show selectivity to virtual depth of the stimuli which could be estimated from motion parallax. In other words, V1 neuronal responses during active locomotion can be characterised as 3D receptive fields. This result provides a revolutionary perspective of how visual scenes are parsed by V1. V1 contains a 3D map of the visual scene, with each neuron representing

a 3D location of the environment.

Similar strategies of encoding the visual scene have been applied in computer vision. Depth information can be represented in the form of a depth map, where each pixel on a 2D image contains the distance from the camera to the corresponding point in the visual scene (Müller et al., 2011). Depth maps can be generated by using cameras in the way of mimicking the depth perception cues used by animals, such as stereo matching (analogous to binocular disparity), structure from motion (motion parallax), structure from shading and depth from defocus (pictorial cues) as well as measuring the time of reflection of light (similar to a sonar system) (Ding & Sharma, 2017; Ibrahim et al., 2020; Li et al., 2006; Waters & Abulula, 2007; Yu & Gallup, 2014). Depth maps can be easily combined with RGB images to facilitate various visual processing tasks such as 3D object recognition and segmentation (Cheng et al., 2023; Jordan, 2013; Ma et al., 2021; Weng & Kitani, 2019) and 3D reconstruction (Sayed et al., 2022; Ylimäki et al., 2018). The understanding of the 3D structure of the visual scenes enables the computer system to interact with our 3D world and powers technologies such as augmented reality (AR, Du et al., 2020; Gimeno et al., 2013 and autonomous driving Fu et al., 2019; Liao et al., 2020; You et al., 2020. Studying how the depth is parsed by biological agents may contribute to advancing the algorithms for 3D artificial agents. For example, understanding how the brain integrates multiple depth cues to converge on a consistent perception of depth may facilitate decision making in autonomous agents (such as robots or self-driving cars) operating in dynamic and complex environments.

7.3.3 A comprehensive depth map in the visual cortex

In this project, I focused on motion parallax and demonstrated that V1 forms a depth map of the visual space from motion parallax when actively navigating in a VR environment. Individual V1 neurons encode specific three-dimensional locations of the visual cues, with different neuronal populations representing near and far visual cues. Although neurons with distinct depth preferences are largely intermingled with each other, a correlation between depth preferences and the anterior-posterior positions of individual neurons was observed. Near depths are over-represented in the upper and lateral visual field, while far-preferring neurons are more common in the lower visual field. Apart from motion parallax, visual scenes in the real world usually consist of rich binocular and monocular depth cues. The visual cortex may create a comprehensive depth map by integrating various depth cues to facilitate a coherent depth perception.

For mice, the main binocular depth cue is binocular disparity (Samonds et al., 2019). Depth from binocular disparity is represented by the binocular zone of V1 as well as HVAs such as RL and LM (La Chioma et al., 2019, 2020; Scholl et al., 2013). Interestingly, similar to the distribution of depth representation from motion parallax, disparity preferences in binocular V1 are largely intermingled (La Chioma et al., 2020). However, fine-scale clustering of neurons with similar disparity preferences has been observed (La

Chioma et al., 2020). Moreover, depth preferences from binocular disparity are related to the elevation of RFs of neurons. V1 neurons with RFs in the lower half of the visual field prefer nearer depths compared to V1 neurons representing the upper visual field (La Chioma et al., 2019). RL, mainly covering the lower visual field, also prefer nearer depths compared to V1. This is opposite to what I observed in V1 for depth tuning from motion parallax – the part of V1 covering the upper visual field prefers near depths. This distinction may stem from the difference in functional significance of the binocular and monocular zones in mice. Mice use the binocular zone to support hunting behaviour while keeping the prey in a retinal area with minimal optic flow (Holmgren et al., 2021; Johnson et al., 2021). The lower visual field of the binocular zone may be primarily used for locating prey during hunting. On the other hand, the upper visual field of the wide monocular zone may specialise in the detection of nearby predators, when visual motion is crucial and binocular information may not be always available.

The majority of pictorial cues, including perspectives, relative size, texture gradients and shading, requires integrating and comparing visual features from multiple retinotopic locations. In humans and primates, area IT, part of the ventral stream, was shown to be involved in inferring depth from pictorial cues such as shading, perspective and texture gradients (Heider, 2000; Liu et al., 2004; Marotta et al., 1997; Mon-Williams et al., 2001; Turnbull et al., 2004). In mouse, texture was found to be represented mainly by LM and LI, analogous to the ventral stream areas in primates for object recognition (Yu et al., 2022). Existing evidence on the use of pictorial cues to aid depth perception has been highly limited in mice. If mice do make depth judgements based on pictorial cues, it may be supported by HVAs which integrate inputs from V1 and other HVAs conveying local visual features such as spatiotemporal frequency, orientation, size and luminance (Adesnik et al., 2012; Andermann et al., 2011; Bonin et al., 2011; Dipoppa et al., 2018; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008).

The same cells in the visual cortex may receive multiple depth cues, such as disparity and motion parallax signals, at the same time. How can they integrate all depth cues to form a coherence perception of depth? Behavioural studies in humans show that disparity and motion parallax are not processed independently by the human visual system (Bradshaw & Rogers, 1996; Domini et al., 2006). When the depth information provided by disparity and motion parallax was conflicted, the perceived depth was better predicted by a composite score between disparity and motion parallax signals rather than a weighted linear summation of both signals (Domini et al., 2006). Studies in macaques have found that MT neurons either show matched or opposite preferences of depth sign in response to motion parallax vs. binocular disparity cues (congruent cells vs. opposite cells, Nadler et al., 2013). When depth was specified by both binocular disparity and motion parallax cues, depth selectivity was enhanced in congruent cells but was slightly reduced in opposite cells. Therefore, congruent cells may be responsible for integration of depth cues for more accurate depth perception. On the other hand, opposite cells may signal the conflict

between motion parallax and disparity cues (Nadler et al., 2013), which contributes to the detection of object motion during self-motion (Kim et al., 2022; Kim et al., 2016; Nadler et al., 2013). It will be interesting to measure the neuronal responses to motion parallax and binocular disparity in the binocular zone of V1 or HVAs in mice to see how neurons may support depth cue integration and object motion detection.

7.3.4 The role of V1 in depth perception from motion parallax across species

It is crucial to recognize that visual cortex can play distinct roles in depth perception for rodents and for other mammals such as cats and primates. Although the visual cortex is indispensable for depth perception in rodents including mice, rats, and Mongolian gerbils (Ellard et al., 1986; Meyer et al., 1966; Parker, Abe, Beatie, et al., 2022), lesions in the cat neocortex or just the visual cortex did not affect the performance of cats on the visual cliff (Cornwell et al., 1976; Meyer et al., 1966; Meyer, 1963), suggesting that the visual cortex is not required for cats' depth perception. Rodents share many similarities in the functional organisations of the visual cortex with carnivores and primates, while also displaying notable differences. For example, rodent visual cortex lack the clear columnar structures in higher mammals for orientation and spatial frequency selectivity (Hubel & Wiesel, 1962, 1963; Hubel & Wiesel, 1968, 1972, 1974; Hubel et al., 1977; Issa et al., 2000; Ohki & Reid, 2007; Ohki et al., 2005) (although clusters of similarly-tuned cells exist, see Ringach et al., 2016). Neurons in mouse V1 have much more complicated receptive fields compared to Gabor-like receptive fields in macaques (Fu et al., 2024). Moreover, their V1 activity is modulated by locomotion to different degrees – mouse V1 neurons are widely modulated by locomotion (Keller et al., 2012; Niell & Stryker, 2010; Saleem et al., 2013), whilst spontaneous facial and body movement in macaques produce minimal effect on their V1 neuronal activity (Galletti et al., 1984; Talluri et al., 2023). Therefore, while V1 neurons in mice play a more significant role in processing motion parallax cues to support depth perception, depth computation from motion parallax may predominantly happen in HVAs for carnivores and primates. For example, many studies have identified the area MT as a crucial structure supporting depth computation from motion parallax for primates (Kim et al., 2015a; Kim et al., 2016; Nadler et al., 2008, 2009).

7.4 Functions of locomotion-related modulation in V1

Locomotion-related modulation of V1 activity has been extensively characterised (Ayaz et al., 2013; Dadarlat & Stryker, 2017; Eriskien et al., 2014; Keller et al., 2012; Niell & Stryker, 2010; Stringer et al., 2019) and the function of such modulation in visual processing has been widely debated (Christensen & Pillow, 2022; Dadarlat & Stryker, 2017;

Keller & Mrsic-Flogel, 2018; Keller et al., 2012; Muzzu & Saleem, 2021; Niell & Stryker, 2010). Some suggested that locomotion modulates the gain of sensory responses in V1, thus enhancing the encoding of visual information in V1 populations (Bennett et al., 2013; Dadarlat & Stryker, 2017; Harris & Thiele, 2011; Niell & Stryker, 2010). Others proposed that visuomotor integration in V1 signals the prediction error between predicted and actual sensory feedback to support predictive coding and sensory learning (Jordan & Keller, 2020; Keller & Mrsic-Flogel, 2018; Keller et al., 2012; Muzzu & Saleem, 2021; Zmarz & Keller, 2016). In this section, I will compare the leading hypotheses on the functional significance of locomotion-related modulation in V1, and argue that visuomotor integration enables V1 neurons to facilitate depth estimation from motion parallax.

7.4.1 Cortical state modulation and enhanced encoding of visual information

It has been proposed that locomotion indiscriminately increases visual responses in V1 which amplifies visual signals (Niell & Stryker, 2010). Locomotion, similar to attention or arousal (Harris & Thiele, 2011; Iriki et al., 1996; Lee & Dan, 2012; Reimer et al., 2014; Vinck et al., 2015), was thought to induce a general modulation of cortical state that elevates the gain of sensory responses in V1 (Bennett et al., 2013; Harris & Thiele, 2011; Niell & Stryker, 2010). Niell and Stryker (2010) found that when mice were running while viewing grating stimuli, there was a sharp elevation of visually-evoked responses when the running speed exceeded 1 cm/s, but no further rise in activity was observed as speed increased. This increased gain of sensory responses did not change the selectivity towards sensory stimuli such as orientation tuning (Niell & Stryker, 2010). Locomotion was found to correlate with reduced membrane potential variability in V1 neurons by suppressing spontaneous firing rate (Bennett et al., 2013). Meanwhile it boosted the sub-threshold responses to visual stimulation (Bennett et al., 2013). Therefore, it was proposed that locomotion imposes a generic modulation of V1 activity by changing the cortical state broadly in a nearly binary manner (Niell & Stryker, 2010).

The modulation on cortical state can contribute to enhancing encoding of visual information and assist visually-guided tasks. When presenting mice on a treadmill with drifting gratings moving at multiple directions, running increased the mutual information between neuronal responses in all V1 layers and visual stimuli, and the parameters of visual stimuli could be more accurately decoded from V1 population activity during running compared to stationary periods (Dadarlat & Stryker, 2017). Running also significantly reduces the noise correlations between V1 neuronal pairs, allowing more robust neural representation of visual stimuli across trials (Dadarlat & Stryker, 2017; Erisken et al., 2014). A more accurate encoding of visual information can lead to better performance for visually-guided tasks. When testing trained mice on a visual detection task for contrast discrimination of gratings, locomotion significantly promoted success discrimination rate (Bennett et al., 2013). Nevertheless, the effect of locomotion on visually-guided tasks

may vary with behavioural context. For example, when mice were trained on a task to detect changes in grating contrast, running significantly impaired the performance of mice trained on a spatially-selective task but not the non-selective group (McBride et al., 2019). This suggests that locomotion may compete with mechanisms supporting selective attention when locomotion is not relevant to the task (McBride et al., 2019).

On the other hand, our data and some other previous studies (Christensen & Pillow, 2022; Saleem et al., 2013) do not support that locomotion only induces a generic change in cortical state. Instead, locomotion-related modulation in V1 operates in a more complex manner. In our datasets, running did not modulate the activity of all V1 neurons in the same manner – V1 populations cover a wide range of running speed preferences, and running speed signals modulate the gain of optic flow responses with the greatest responses at the conjunction of individual neuron’s preferred running speed and optic flow speed. This modulation is neither binary nor monotonic, as the degree of modulation varies continuously with running speed, and the amplitude of modulation does not merely monotonically increase or decrease. Non-monotonic running speed tuning in V1 neurons has been reported before, when V1 neuronal activity was measured in the dark or during visual stimulation while mice were running on a treadmill (Christensen & Pillow, 2022; Saleem et al., 2013). Therefore, locomotion does not merely induce a simple generic or binary change to sensory responses in V1, but modulates neuronal activity according to individual neuron’s running speed tuning. Nevertheless, it is possible that locomotion does change the cortical state and applies an uniform effect on all neurons to some degree, whereas running speed tuning in individual neurons provides further activation of specific populations related to the visual stimuli or tasks. It is also possible that some neurons are affected by running in a binary manner, whereas others have specific selectivity for running speed.

7.4.2 Predictive coding

Another hypothesis on the function of locomotion-related modulation pertains to predictive processing (Jordan & Keller, 2020; Keller & Mrsic-Flogel, 2018; Keller et al., 2012; Muzzu & Saleem, 2021; Zmarz & Keller, 2016). There are two main streams of thoughts regarding how neural circuits encode information (Keller & Mrsic-Flogel, 2018) – the representational framework (Marr, 2010) and the predictive processing framework (Clark, 2016). The representation framework suggests that neuronal activity in the sensory system is mostly driven by bottom-up feed-forward inputs, with layers of feature detectors representing different levels of object features in the environment (e.g. from edges to shapes and faces, Martin, 1994). On the other hand, the predictive processing framework proposes that the brain can form and use an internal model of the world to make predictions about sensory inputs based on current locomotion and past sensory experience (Friston, 2005; Körding & Wolpert, 2004; Rao & Ballard, 1999; Spratling, 2010). In this framework, cortical sensory responses are driven by the prediction errors between

the actual bottom-up sensory signals and the top-down prediction of sensory inputs based on a generative model of the world. In this way, the internal representation can be updated based on prediction errors or even in absence of bottom-up inputs, which helps sensorimotor learning and simulating the environment (Keller & Mrsic-Flogel, 2018).

The idea that predictive coding underlies locomotion-related modulation in V1 came from a series of experiments when visuomotor mismatch was created in the visual stimuli (Jordan & Keller, 2020; Keller & Mrsic-Flogel, 2018; Keller et al., 2012; Muzzu & Saleem, 2021; Zmarz & Keller, 2016). A small subset of L2/3 V1 neurons (13% measured by Keller et al., 2012) were found to be responsive (or displaying subthreshold responses) to a mismatch between running speed and visual flow speed created by a sudden halt in visual flow (mismatch neurons), but not responding to passive viewing of visual flow halts (Jordan & Keller, 2020; Keller et al., 2012; Zmarz & Keller, 2016). This was interpreted as prediction error responses to the mismatch between bottom-up optic flow speed input and top-down predictions of optic flow speed based on current running speed.

However, as discussed in Chapter 5, the depth-selective responses observed in V1 seem not to be a result of prediction error responses. First, since optic flow speed is determined by the ratio of running speed and depth, optic flow speed cannot be accurately predicted without the information on depth. Second, during closed loop experiments when visuomotor mismatch was absent, depth-selective neurons were active as long as stimuli were presented at their preferred depths and retinotopic locations. Third, the conjunctive coding of running speed and optic flow speed does not depend on closed loop coupling of locomotion and visual feedback. During open loop experiments when the predicted optic flow did not match the actual sensory feedback, depth-selective neurons displayed similar peak response amplitudes compared to closed loop conditions. The optic flow speed preferences and running speed preferences of individual neurons were also maintained in open loop. Therefore, depth selectivity in V1 observed in my experiments does not reflect prediction error.

Mismatch responses described in previous studies can be explained by depth selectivity. Neurons selective for negative prediction errors (slower visual flow than predicted) may reflect a preference for a lower gain of optic flow responses and correspond to far-preferring cells, whereas neurons selective for positive prediction errors (faster visual flow than predicted) may be a subset of near-preferring cells. Alternatively, visuomotor prediction error and depth may be represented by independent neuronal populations in V1. While mismatch neurons identified in previous studies only occupy a small proportion of V1 neurons, depth-selectivity seems more widespread. Another possibility is that depth-selective responses in V1 neurons may form the basis for predictive coding in downstream areas. The prediction of optic flow can be refined with depth information, contributing to a more accurate internal model of the world.

7.5 The role of V1 depth selectivity in spatial perception

Mouse V1 is largely organised in retinocentric coordinates which were inherited from early visual areas (Dräger, 1975; Kalatsky & Stryker, 2003; McLaughlin & O’Leary, 2005; Wagor et al., 1980). In order to perform goal-directed behaviour such as navigation towards a target or prey capture, the environment needs to be represented in an egocentric coordinate system, so that the environment is anchored to the position of a given body part and the representation remains stable regardless of eye movements (Martins et al., 2024). Egocentric information can then be used to form an allocentric representation of the environment, which is invariant to the animal’s pose and facilitates the formation of an internal model of the world for spatial navigation (Wang, Chen, & Knierim, 2020).

My experiments leveraged the fact that head-fixed mice make far less eye movements compared to freely-moving mice (Holmgren et al., 2021; Meyer et al., 2020; Michaiel et al., 2020). They do not exhibit non-conjugate eye movements compensatory to head tilts for visual field stabilisation, but only make conjugate eye movements in a "saccade and fixation" manner in response to head rotation attempts. The analysis of eye movements of mice in my experiments also confirmed that eye positions were generally stable during the recording and saccades were rare (Figure 4.8I-L). Therefore, I did not calibrate the distance of sphere stimuli according to eye movements. However, in freely-moving conditions, the activity of depth-selective neurons in V1 are expected to anchor to the retinal image, and gaze direction changes caused by eye and head movements would drastically affect neuronal responses in V1 even if the animal’s body remains stationary with respect to the stimuli.

Therefore, to compute the location of stimuli in an egocentric coordinate (i.e. with respect to the animal’s body and invariant to gaze direction), V1 depth-selective responses or depth signals from other sources in retinocentric coordinates needs to be integrated with gaze directions to form an egocentric representation of object locations in the environment. A theoretical model for the process of transforming retinocentric to egocentric representation has been proposed (Martins et al., 2024; Pouget, 1993; Zipser and Andersen, 1988, Figure 7.1). First, visual inputs in retinotopic coordinates and eye position signals are integrated in neurons that jointly encode both visual stimuli and eye position. Next, the activity of joint-coding neurons are combined which gives rise to egocentric representation that does not vary with eye position. This is similar to the model discussed in Chapter 5 (Figure 5.11) of how depth coding can arise from joint coding of running and optic flow speeds.

In mammals, information on eye position can be obtained from proprioceptive signals indicating extraocular muscle position and the efference copies of the eye movement motor command (Ambrad Giovannetti & Rancz, 2024; Wurtz, 2008). Motor command that directly innervate in the eye muscles in mice comes from the oculomotor nuclei in

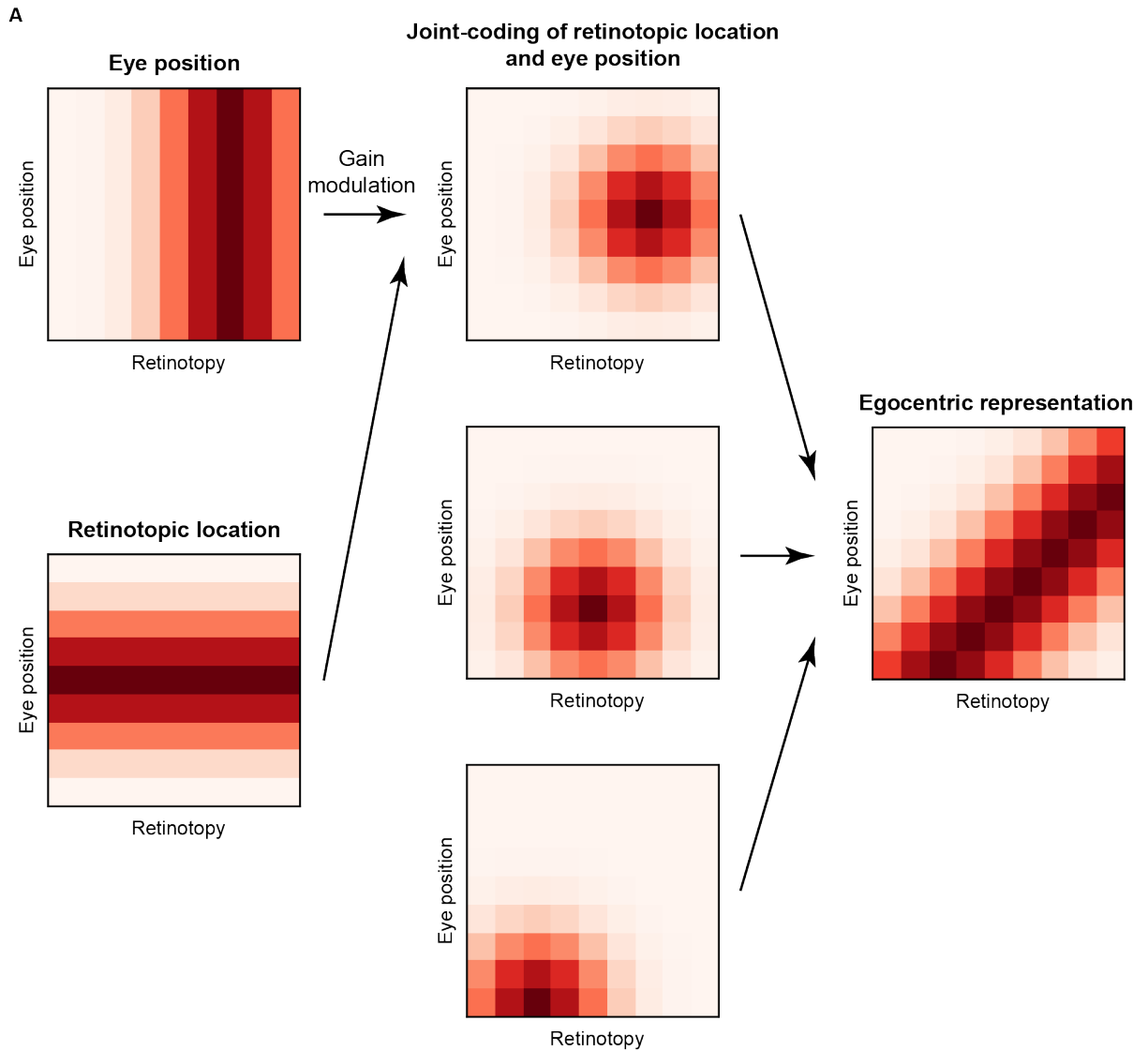


Figure 7.1: Model of the transformation from retinocentric to egocentric representation. (A) Model schema of the transformation from retinocentric to egocentric representation, adapted from Figure 5.11 and Martins et al., 2024. At the early stage of visual processing, visual inputs are encoded in retinocentric coordinates (left). At intermediate stages, retinotopic inputs and eye position signals are integrated in neurons that jointly encode both visual stimuli and eye position (middle). Next, inputs from joint-coding neurons are combined, giving rise to egocentric representation that is invariant to eye position (right).

the midbrain and brainstem, which is controlled by multiple subcortical and cortical areas such as the accessory optic system, the vestibular system, cerebellum, SC, RSC, the thalamus, the visual cortex and the motor cortex (Ambrad Giovannetti & Rancz, 2024; Sparks, 2002). In mice, eye positions are strongly coupled with head movements (Holmgren et al., 2021; Meyer et al., 2020; Michael et al., 2020). The visual cortex can obtain head movement signals from vestibular organs (Cullen, 2019; Vélez-Fort et al., 2018), SC (Wang et al., 2015; Wilson et al., 2018) and secondary motor cortex (M2, Guitchounts et al., 2020).

Although neurons that jointly code depth and gaze direction have not been directly identified, the modulation of eye or head movements of neuronal activity has been identified across species. Surprisingly in mice, encoding of head and eye movements can be found in V1 in addition to the representation of visual information (Bouvier et al., 2020; Guitchounts et al., 2020, 2022; Meyer et al., 2018; Parker, Abe, Leonard, et al., 2022). Bouvier et al. (2020) explored the modulation of V1 activity by simulated head movement induced by body rotation in the dark and light conditions without explicit visual stimuli. They found that while head movement mainly suppressed V1 activity in the dark, head movement excited V1 neurons across all cortical layers in ambient light. Guitchounts et al. (2020) further investigated in the encoding of head-orienting movements in V1 in freely-moving rats. Similar to Bouvier et al. (2020), head-orienting movements were correlated with suppression of overall V1 activity in the dark and elevation in V1 activity in the light. Moreover, there are V1 cells tuned to direction of head movements, yet their head direction tuning was not correlated between the dark and light conditions. The studies above show that V1 activity is modulated by head movements, but the conclusion can be confounded by the changes in retinal images as a result of gaze changes. Using miniature head-mounted cameras (Meyer et al., 2018; Parker, Abe, Leonard, et al., 2022) allows the real-time tracking of the eye and head positions of the animal to obtain the gaze-corrected, retinocentric visual scene. Parker, Abe, Leonard, et al. (2022) show that the activity of many V1 neurons is best described by a multiplicative integration between visual inputs and eye and head positions. This again challenges the traditional view of V1 being a truthful representation of the retinal image. Modulation of neuronal activity by eye or head positions or gaze directions have been found in V1, V2, V3, V4, MT, MST, V6, and PPC of primates (Andersen & Mountcastle, 1982; Andersen et al., 1985; Bremmer et al., 1997; Durand et al., 2010; Galletti & Battaglini, 1989; Galletti et al., 1995; Guo & Li, 1997; Morris & Krekelberg, 2019; Morris et al., 2012; Rosenbluth & Allman, 2002; Sakata et al., 1980; Trotter & Celebrini, 1999) and the visual cortex of humans (Merriam et al., 2013).

Finally, egocentric representation of distances can be found in various areas. In areas such as retrosplenial cortex (RSC), dorsomedial striatum (DMS), lateral entorhinal cortex (LEC), postrhinal cortex (POR) and PPC, some neurons were found to be excited or inhibited when rats were at a specific distance to any wall and are moving towards the wall in a specific heading direction or body orientation (egocentric boundary vector cells), and their responses to local boundaries can be generated across environments (Alexander et al., 2020, 2022; Gofman et al., 2019; Hinman et al., 2019; Wang et al., 2018). Some LEC cells are also tuned to distance to 3D objects or goal at a specific bearing (Wang et al., 2018). It was proposed that these egocentric spatial representation can be combined with heading direction and body movements to generate allocentric spatial representation in median entorhinal cortex (MEC) and the hippocampal formation (Bicanski & Burgess, 2018; Byrne et al., 2007; Hinman et al., 2019; Martins et al., 2024). Examples of allocentric

coding of distances include allocentric boundary vector cells found in the subiculum of the hippocampal formation, which would fire when the animal was at the specific distance to a boundary in a specific allocentric direction (regardless of the animal's head or body direction, Lever et al., 2009). The firing fields of boundary vector cells remained stable across environments with boundaries of different colours, shapes, or materials and across environments with or without vertical walls (Lever et al., 2009). Border cells previously found in MEC that fire along environmental boundaries (Savelli et al., 2008; Solstad et al., 2008) might include a subset of boundary vector cells that are tuned to short distances (Lever et al., 2009). Boundary vector cells have been proposed to provide environmental inputs to the hippocampus to assist the formation of place fields in place cells (Burgess et al., 2000; O'Keefe & Burgess, 1996) that represent specific allocentric spatial locations and are thought to be key for forming a cognitive map of the space (O'Keefe, 1976). Similarly, cells that encode the distance and the animal's allocentric direction to discrete objects in the environment (object vector cells) were also identified in the MEC (Høydal et al., 2019). It is also proposed that allocentric spatial information can be converted back to egocentric representation to facilitate behavioural output (Bicanski & Burgess, 2018; Byrne et al., 2007; Hinman et al., 2019).

In summary, depth-selective responses in V1 can form an integral part of the wider spatial navigation system. Combined with signals conveying eye and head positions and body movements, the representation of depth in retinotopic coordinates in V1 can be used to create egocentric and allocentric representation of three-dimensional locations of objects in the environment, which enables animals to perform spatial navigation, prey capture and other visually-guided tasks.

7.6 Circuit mechanisms of depth coding in mouse V1

7.6.1 Long-range inputs of locomotion signals

Depth-selective responses in V1 neurons require the integration between visual motion and locomotion-related signals. This integration might be inherited from upstream subcortical inputs to V1 or generated in V1 *de novo*. For example, some degrees of locomotion-related modulation can already be observed in dLGN, LP and SC. Although Niell and Stryker (2010) did not find a significant change in visually-driven LGN activity during running compared to stationary periods, Erisken et al. (2014) later found that both spontaneous and visually-driven neuronal activity in dLGN were enhanced at locomotion onset, although to a smaller degree compared to V1. When imaging neuronal responses to grating stimuli in a running mice under an open-loop condition, (Roth et al., 2016) found that $\sim 20\%$ dLGN boutons in V1 exhibited activity that cooperatively integrate running and visual motion speed signals. On the other hand, many of the LP boutons displayed anticorrelated tuning to running and visual motion speed, suggesting that they may signal the visuomotor mismatch between predicted visual motion speed

based on running speed and actual visual motion speed (Roth et al., 2016). In superficial SC, running did not exert a prominent trend (increase or decrease) in the activity of the whole population (Savier et al., 2019). However, there was a slight increase in activity of a proportion of cells (20%) and slight decrease in some other cells (10%) during running compared to stationary periods (Savier et al., 2019). However, this modulation in SC was much weaker than what was observed in V1, with smaller spontaneous firing rate and response variability (Savier et al., 2019). Therefore, although V1 can potentially inherit some locomotion-related modulation of visual responses from subcortical inputs, a greater degree of visuomotor integration is likely to happen within V1.

V1 can also directly receive running speed signals. The basal forebrain receives projections from the mesencephalic locomotor region which elicits and maintains locomotion (Dautan et al., 2021; Lee et al., 2014; Nauta & Kuypers, 1958; Noga & Whelan, 2022; Shik & Orlovsky, 1976). During locomotion, cholinergic innervations from the basal forebrain directly activate the nicotinic receptors of VIP neurons in V1 (Fu et al., 2014). The secondary motor cortex (M2) and the A42b region of anterior cingulate cortex also provide a major input to L2/3 and L6 excitatory cells and PV interneurons in V1 (Leinweber et al., 2017).

To further investigate how long-range inputs contribute to locomotion signals in V1, inactivation of key input areas can be performed to examine the effects on V1 depth-selective responses.

7.6.2 Local circuitry for visuomotor integration

Local excitatory-inhibitory circuitry within V1 may provide a basis for multiplicative visuomotor integration in V1 pyramidal neurons. Previous research has found that VIP neurons, a group of interneurons that mainly inhibit SST inhibitory neurons (Jiang et al., 2015; Pfeffer et al., 2013), elevated their responses during running with or without visual stimulation (Fu et al., 2014; Jackson et al., 2016; Reimer et al., 2014). On the other hand, in darkness, the activity of SST cells were suppressed by locomotion, and PV populations showed both positive and negative association with locomotion (Fu et al., 2014). Therefore, it was proposed that locomotion signals activate VIP neurons and elevate the activity of pyramidal neurons via disinhibition (Fu et al., 2014).

Apart from a broad increase in visual responses, can disinhibition contribute to the conjunctive coding of running speed and optic flow speed that I observed in Chapter 5? Conjunctive coding requires a multiplicative integration of running speed and optic flow speed tuning in individual neurons. Disinhibition has been implicated in multiplicative computation in T4 neurons, supporting the computation of direction of visual motion in *Drosophila melanogaster* (Groschner et al., 2022). The supralinear integration of signals for visual motion computation in individual T4 direction-selective neurons comes from the coincidence of cholinergic excitation and the release from glutamatergic shunting inhibition through the reduction in conductance (Groschner et al., 2022). The multiplicative

integration of running and optic flow speed could be a result of a similar disinhibitory circuitry, where excitatory inputs from optic flow speed-tuned neurons coincides with the disinhibitory signals carrying running information from VIP neurons. This mechanism would rely on the formation of specific connections between excitatory and inhibitory neuronal pairs with opposite running speed tuning.

On the contrary, evidence have been found to challenge the disinhibition model of gain modulation in V1. Polack et al. (2013) using in vivo current-clamp whole-cell recordings has found an increase in both spontaneous and visually evoked firing rates in both PV neurons and SST neurons during running (although only 9 PV neurons and 10 SST neurons were sampled due to restrictions of the recording technique). Pakan et al. (2016) further suggested that locomotion-related modulation of neuronal activity depends on the visual context. Locomotion increased the activity of VIP neurons and PV neurons during running both in darkness and during visual stimulation, and visual stimulation slightly reduces the average activity of VIP neurons. On the other hand, SST neuronal responses to locomotion were more context-dependent. During visual stimulation, locomotion significantly increased SST neuronal responses. In darkness, SST neurons could increase or decrease their activity or stay non-responsive during locomotion, and the SST population on average show no consistent changes in responses during locomotion. Overall, visual stimulation significant increased the average responses of SST neurons as well as the correlation between running speed and neuronal responses. Therefore, Pakan et al. (2016) proposed that disinhibition may explain the activity of a small proportion of SST neurons whose activity was inhibited during locomotion in the darkness. On the contrary, locomotion-related modulation in the rest of neurons may manifest in a more context-dependent way. Neuromodulatory inputs evoked by locomotion may not only activate VIP neurons, but also directly activate SST, PV and excitatory neurons in V1. In darkness, SST neurons largely remain unresponsive to locomotion due to sub-threshold influence of neuromodulatory inputs and VIP neuron activation. During visual stimulation, both SST and excitatory neurons are activated by overcoming the inhibition by VIP and PV neurons.

A recurrent network model similar to what was proposed by (Salinas & Abbott, 1996) may be an alternative way of explaining the multiplicative integration of running speed and optic flow speed that gives rise to depth selectivity. Neurons are recurrently connected in the network of the visual cortex, with potential connections forming between excitatory neurons with similar running speed and optic flow speed preferences. In this way, excitatory external inputs carrying similar running speed and optic flow speed information can be amplified in groups of recurrently connected neurons with like-to-like connections in a multiplicative manner. On the other hand, inhibitory neurons, which are often broadly tuned (Packer & Yuste, 2011; Xue et al., 2014; Znamenskiy et al., 2024), may form broader connections with excitatory neurons with slightly different depth preferences, regulating the network dynamics. Dipoppa et al. (2018) further constructed a

recurrent network model for predicting visuomotor responses in different cell types in V1. In this model, locomotion increases external excitatory visual inputs to all cell types while decreasing recurrent excitation between pyramidal cells and boosting recurrent inhibition between PV cells. The effects of locomotion on visual responses depends on the size of the stimuli, with VIP cell activity being modulated by small stimuli and SST cells by large stimuli.

Nevertheless, the studies above usually treat running as a binary state – stationary or locomotion (Fu et al., 2014; Jackson et al., 2016; Jiang et al., 2015; Pakan et al., 2016; Pfeffer et al., 2013; Polack et al., 2013; Reimer et al., 2014), whereas neuronal responses to running in V1 is largely dependent on the running speed and the running speed tuning of individual neurons (shown in Chapter 5, Christensen and Pillow, 2022; Saleem et al., 2013). It remains unclear how the connections and dynamic interactions between different cell types give rise to the multiplicative visuomotor integration observed in this study. As we are still discovering new excitatory-inhibitory connection motifs (Kuan et al., 2024; Znamenskiy et al., 2024), further investigation is needed to fully uncover the connection patterns between different types of excitatory and inhibitory neurons, and to investigate how local circuitry contribute to visuomotor integration in V1.

7.7 Conclusion and future outlooks

In conclusion, my project serves as a crucial step toward advancing our understanding of neural mechanisms underlying depth perception – an essential part of animal behaviour in this 3D world. Using mice as the model organism, I first established the importance of depth perception in mice using visual cliff as a behavioural assay and show that depth perception is innate for mice. Then, I focused on motion parallax, a powerful monocular depth cue for mice which remained largely unexplored. Combining two-photon calcium imaging and 3D VR environments, I show that pyramidal neurons in V1 have three-dimensional receptive fields during active locomotion – they encode not only the retinotopic locations but also virtual depths of the visual stimuli. The representation of depth is widespread in V1, with neurons representing similar retinotopic location being able to respond to a wide range of depths. Depth selectivity from motion parallax arises from the conjunctive coding of running speed and optic flow speed in a multiplicative manner. Overall, this study offers a groundbreaking perspective on the role of V1 activity in visual processing.

This project opens up many exciting avenues for future research. One important direction is to characterise depth perception (and visual responses in general) of mice in more naturalistic environments. Visual neuroscience has transitioned from using simple parametric stimuli (Andermann et al., 2011; Bonin et al., 2011; Dräger, 1975; Gao et al., 2010; LeDue et al., 2012; Niell & Stryker, 2008) to more naturalistic ones (Bashivan et al., 2019; Fu et al., 2024; Walker et al., 2019) to probe the activity of the visual cortex. This project further emphasizes the importance of incorporating the third dimension –

depth – into visual stimuli. In future experiments, more elements from the real world can be included to make the visual responses studied in lab as naturalistic as possible. For example, mice could be tested in freely moving conditions, with natural scenes featuring richer depth cues, such as binocular and pictorial cues, objects at varying depths, and introducing object motion. VR provides a perfect tool for creating complex visual environments that can be easily manipulated.

Another critical avenue of investigation is to establish a causal link between neuronal activity in the visual system and depth perception. In this project, I measured V1 neuronal responses to stimuli at different distances to the observer. This does not imply a causal relationship between neuronal activity in the visual cortex and depth perception. Parker, Abe, Beatie, et al. (2022) have shown that V1 is required for both binocular and monocular depth perception, yet the role of higher visual areas is unknown. Further experiments are needed to establish a causal link between neuronal activity at different parts of the visual cortex and depth perception, by testing the performance of mice on the visual cliff under optogenetic or chemogenetic inhibition of different visual areas. Once the areas causally involved in depth perception are identified, further research can focus on the role of specific neuronal populations within these regions in depth perception.

A third direction is to systematically characterise the representation of depth across the visual system. In this project, I focused on the depth representation in L2/3 of V1. Further experiments can focus on characterising the depth-selective responses across different subcortical regions, HVAs and layers of the visual cortex. Among HVAs, anterior HVAs (e.g. AM, A, RL) may prefer near depths while posterior HVAs (e.g. PM) may prefer far depths based on their speed preferences (Andermann et al., 2011; Garrett et al., 2014; Glickfeld & Olsen, 2017; Marshel et al., 2011; Roth et al., 2012). Neurons in the shallower and deeper layers of V1 may have different tuning properties to running speed, optic flow speed and depth. In the output layers of V1 or in HVAs, it may be possible to identify neurons that are tuned to depth invariant of running speed or optic flow speed tuning.

Finally, the innate nature of depth perception provides a perfect platform to explore the relationship between gene expression patterns, circuit mechanisms and behavioural functions. In vivo imaging can be combined with high-throughput spatial transcriptomics to link in vivo with ex vivo data. Spatial transcriptomics enable gene expression patterns to be extracted while retaining the spatial information of cells. By registering in vivo imaging data with ex vivo spatial transcriptomics data, functions, connectivity and cell types of the same cells can be connected. Several interesting questions can be explored. First, do different excitatory cell types in V1 vary in their depth preferences? Do inhibitory cell types have depth selectivity? If so, how are they tuned to running speed, optic flow speed and depth? If HVAs specialise in their depth preferences, does this specialisation come from different populations of V1 projection neurons? Exploring these questions will provide valuable insights into the circuit mechanisms underlying depth perception.

Bibliography

- Adams, D. L., & Horton, J. C. (2003). A Precise Retinotopic Map of Primate Striate Cortex Generated from the Representation of Angioscotomas. *The Journal of Neuroscience*, 23(9), 3771–3789. <https://doi.org/10.1523/JNEUROSCI.23-09-03771.2003>
- Adams, D. L., & Horton, J. C. (2009). Ocular Dominance Columns: Enigmas and Challenges. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, 15(1), 62–77. <https://doi.org/10.1177/1073858408327806>
- Adams, D. L., & Zeki, S. (2001). Functional Organization of Macaque V3 for Stereoscopic Depth [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 86(5), 2195–2203. <https://doi.org/10.1152/jn.2001.86.5.2195>
- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America. A, Optics and Image Science*, 2(2), 284–299. <https://doi.org/10.1364/josaa.2.000284>
- Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z. J., & Scanziani, M. (2012). A neural circuit for spatial summation in visual cortex. *Nature*, 490(7419), 226–231. <https://doi.org/10.1038/nature11526>
- Adesnik, H., & Naka, A. (2018). Cracking the Function of Layers in the Sensory Cortex. *Neuron*, 100(5), 1028–1043. <https://doi.org/10.1016/j.neuron.2018.10.032>
- Alexander, A. S., Carstensen, L. C., Hinman, J. R., Raudies, F., Chapman, G. W., & Hasselmo, M. E. (2020). Egocentric boundary vector tuning of the retrosplenial cortex [Publisher: American Association for the Advancement of Science]. *Science Advances*, 6(8), eaaz2322. <https://doi.org/10.1126/sciadv.aaz2322>
- Alexander, A. S., Tung, J. C., Chapman, G. W., Conner, A. M., Shelley, L. E., Hasselmo, M. E., & Nitz, D. A. (2022). Adaptive integration of self-motion and goals in posterior parietal cortex. *Cell Reports*, 38(10), 110504. <https://doi.org/10.1016/j.celrep.2022.110504>
- Ambrad Giovannetti, E., & Rancz, E. (2024). Behind mouse eyes: The function and control of eye movements in mice. *Neuroscience & Biobehavioral Reviews*, 161, 105671. <https://doi.org/10.1016/j.neubiorev.2024.105671>
- Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L., & Reid, R. C. (2011). Functional specialization of mouse higher visual cortical areas. *Neuron*, 72(6), 1025–1039.

- Andersen, A., & Mountcastle, B. (1982). The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *The Journal of Neuroscience*, *3*, 532–548.
- Andersen, G. J., & Braunstein, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, *34*(4), 356–362. <https://doi.org/10.3758/BF03203048>
- Andersen, R. A., Essick, G. K., & Siegel, R. M. (1985). Encoding of Spatial Location by Posterior Parietal Neurons [Publisher: American Association for the Advancement of Science]. *Science*, *230*(4724), 456–458. <https://doi.org/10.1126/science.4048942>
- Anzai, A., Chowdhury, S. A., & DeAngelis, G. C. (2011). Coding of Stereoscopic Depth Information in Visual Areas V3 and V3A. *Journal of Neuroscience*, *31*(28), 10270–10282. <https://doi.org/10.1523/JNEUROSCI.5956-10.2011>
- Attinger, A., Wang, B., & Keller, G. B. (2017). Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex [Publisher: Elsevier]. *Cell*, *169*(7), 1291–1302.e14. <https://doi.org/10.1016/j.cell.2017.05.023>
- Ayaz, A., Saleem, A. B., Schölvinck, M. L., & Carandini, M. (2013). Locomotion Controls Spatial Integration in Mouse Visual Cortex. *Current Biology*, *23*(10), 890–894. <https://doi.org/10.1016/j.cub.2013.04.012>
- Baden, T., Berens, P., Franke, K., Rosón, M. R., Bethge, M., & Euler, T. (2016). The functional diversity of retinal ganglion cells in the mouse. *Nature*, *529*(7586), 345–350. <https://doi.org/10.1038/nature16468>
- Baker, T. I., & Issa, N. P. (2005). Cortical Maps of Separable Tuning Properties Predict Population Responses to Complex Visual Stimuli [Place: US Publisher: American Physiological Society]. *Journal of Neurophysiology*, *94*(1), 775–787. <https://doi.org/10.1152/jn.01093.2004>
- Bakin, J. S., Nakayama, K., & Gilbert, C. D. (2000). Visual Responses in Monkey Areas V1 and V2 to Three-Dimensional Surface Configurations [Publisher: Society for Neuroscience Section: ARTICLE]. *Journal of Neuroscience*, *20*(21), 8188–8198. <https://doi.org/10.1523/JNEUROSCI.20-21-08188.2000>
- Barlow, H. B. (1953). Summation and inhibition in the frog's retina. *The Journal of Physiology*, *119*(1), 69–88. Retrieved December 24, 2024, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1393035/>
- Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *The Journal of Physiology*, *193*(2), 327–342. <https://doi.org/10.1113/jphysiol.1967.sp008360>
- Bartfeld, E., & Grinvald, A. (1992). Relationships between orientation-preference pinwheels, cytochrome oxidase blobs, and ocular-dominance columns in primate striate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(24), 11905–11909. Retrieved December 26, 2024, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC50666/>

- Bashivan, P., Kar, K., & DiCarlo, J. J. (2019). Neural population control via deep image synthesis [Publisher: American Association for the Advancement of Science]. *Science*, *364*(6439), eaav9436. <https://doi.org/10.1126/science.aav9436>
- Basole, A., White, L. E., & Fitzpatrick, D. (2003). Mapping multiple features in the population response of visual cortex [Publisher: Nature Publishing Group]. *Nature*, *423*(6943), 986–990. <https://doi.org/10.1038/nature01721>
- Baxter, W. T., & Dow, B. M. (1989). Horizontal organization of orientation-sensitive cells in primate visual cortex. *Biological Cybernetics*, *61*(3), 171–182. <https://doi.org/10.1007/BF00198764>
- Beier, C., Bocchero, U., Zhang, Z., Jin, N., Massey, S. C., Ribelayga, C. P., Martemyanov, K., Hattar, S., & Pahlberg, J. (2020, September). Retinal circuits driving a non-image forming visual behavior [Pages: 2020.09.08.288373 Section: New Results]. <https://doi.org/10.1101/2020.09.08.288373>
- Bennett, C., Arroyo, S., & Hestrin, S. (2013). Subthreshold Mechanisms Underlying State-Dependent Modulation of Visual Responses [Publisher: Elsevier]. *Neuron*, *80*(2), 350–357. <https://doi.org/10.1016/j.neuron.2013.08.007>
- Berkeley, G. (1709). *An essay towards a new theory of vision*. Jeremy Pepyat.
- Bhagavatula, P. S., Claudianos, C., Ibbotson, M. R., & Srinivasan, M. V. (2011). Optic Flow Cues Guide Flight in Birds [Publisher: Elsevier]. *Current Biology*, *21*(21), 1794–1799. <https://doi.org/10.1016/j.cub.2011.09.009>
- Bicanski, A., & Burgess, N. (2018). A neural-level model of spatial memory and imagery (L. Colgin & M. J. Frank, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, *7*, e33752. <https://doi.org/10.7554/eLife.33752>
- Bickford, M. E., Zhou, N., Krahe, T. E., Govindaiah, G., & Guido, W. (2015). Retinal and tectal driver-like inputs converge in the shell of the mouse dorsal lateral geniculate nucleus [Publisher: Soc Neuroscience]. *Journal of Neuroscience*, *35*(29), 10523–10534.
- Birch, E. E., Gwiazda, J., & Held, R. (1982). Stereoacuity development for crossed and uncrossed disparities in human infants [Publisher: Elsevier]. *Vision research*, *22*(5), 507–513.
- Bishop, P. O., Henry, G. H., & Smith, C. J. (1971). Binocular interaction fields of single units in the cat striate cortex [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1971.sp009508>]. *The Journal of Physiology*, *216*(1), 39–68. <https://doi.org/10.1113/jphysiol.1971.sp009508>
- Bishop, P., & Pettigrew, J. (1986). Neural mechanisms of binocular vision. *Vision Research*, *26*(9), 1587–1600. [https://doi.org/10.1016/0042-6989\(86\)90177-X](https://doi.org/10.1016/0042-6989(86)90177-X)
- Blakemore, C. (1970). The range and scope of binocular depth discrimination in man [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1970.sp009296>]. *The Journal of Physiology*, *211*(3), 599–622. <https://doi.org/10.1113/jphysiol.1970.sp009296>

- Blasdel, G. G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 12(8), 3139–3161. <https://doi.org/10.1523/JNEUROSCI.12-08-03139.1992>
- Bogin, B. (1997). Evolutionary hypotheses for human childhood. *American Journal of Physical Anthropology*, 104(S25), 63–89. [https://doi.org/10.1002/\(SICI\)1096-8644\(1997\)25+<63::AID-AJPA3>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1096-8644(1997)25+<63::AID-AJPA3>3.0.CO;2-8)
- Bonhoeffer, T., & Grinvald, A. (1991). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns [Publisher: Nature Publishing Group]. *Nature*, 353(6343), 429–431. <https://doi.org/10.1038/353429a0>
- Bonin, V., Histed, M. H., Yurgenson, S., & Reid, R. C. (2011). Local Diversity and Fine-Scale Organization of Receptive Fields in Mouse Visual Cortex [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 31(50), 18506–18521. <https://doi.org/10.1523/JNEUROSCI.2974-11.2011>
- Boone, H. C., Samonds, J. M., Crouse, E. C., Barr, C., Priebe, N. J., & McGee, A. W. (2021). Natural binocular depth discrimination behavior In mice explained by visual cortical activity. *Current Biology*. <https://doi.org/10.1016/j.cub.2021.02.031>
- Bouvier, G., Senzai, Y., & Scanziani, M. (2020). Head Movements Control the Activity of Primary Visual Cortex in a Luminance-Dependent Manner [Publisher: Elsevier]. *Neuron*, 108(3), 500–511.e5. <https://doi.org/10.1016/j.neuron.2020.07.004>
- Bradley, D. C., Qian, N., & Andersen, R. A. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques [Publisher: Nature Publishing Group]. *Nature*, 373(6515), 609–611. <https://doi.org/10.1038/373609a0>
- Bradshaw, M. F., & Rogers, B. J. (1996). The Interaction of Binocular Disparity and Motion Parallax in the Computation of Depth. *Vision Research*, 36(21), 3457–3468. [https://doi.org/10.1016/0042-6989\(96\)00072-7](https://doi.org/10.1016/0042-6989(96)00072-7)
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, 25(11), 120–123. Retrieved November 25, 2024, from <https://elibrary.ru/item.asp?id=4934581>
- Braitenberg, V., & Braitenberg, C. (1979). Geometry of orientation columns in the visual cortex. *Biological Cybernetics*, 33(3), 179–186. <https://doi.org/10.1007/BF00337296>
- Braunstein, M. L. (1968). Motion and texture as sources of slant information [Place: US Publisher: American Psychological Association]. *Journal of Experimental Psychology*, 78(2, Pt.1), 247–253. <https://doi.org/10.1037/h0026269>
- Bredfeldt, C. E., & Ringach, D. L. (2002). Dynamics of Spatial Frequency Tuning in Macaque V1 [Publisher: Society for Neuroscience Section: ARTICLE]. *Journal of Neuroscience*, 22(5), 1976–1984. <https://doi.org/10.1523/JNEUROSCI.22-05-01976.2002>

- 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 [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 77(2), 944–961. <https://doi.org/10.1152/jn.1997.77.2.944>
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Visual Neuroscience*, 10(6), 1157–1169. <https://doi.org/10.1017/s0952523800010269>
- Britten, K. H. (2003, November). The Middle Temporal Area: Motion Processing and the Link to Perception. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (pp. 1203–1216). MIT press. Retrieved October 15, 2024, from https://www.researchgate.net/publication/364581168_The_Middle_Temporal_Area_Motion_Processing_and_the_Link_to_Perception
- Buckthought, A., Yoonessi, A., & Baker, C. L., Jr. (2017). Dynamic perspective cues enhance depth perception from motion parallax. *Journal of Vision*, 17(1), 10. <https://doi.org/10.1167/17.1.10>
- Burgess, N., Jackson, A., Hartley, T., & O’Keefe, J. (2000). Predictions derived from modelling the hippocampal role in navigation. *Biological Cybernetics*, 83(3), 301–312. <https://doi.org/10.1007/s004220000172>
- Burkhalter, A., & Van Essen, D. (1986). Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *The Journal of Neuroscience*, 6(8), 2327–2351. <https://doi.org/10.1523/JNEUROSCI.06-08-02327.1986>
- Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future. *Psychological review*, 114(2), 340–375. <https://doi.org/10.1037/0033-295X.114.2.340>
- Callaway, E. M. (2001). Neural Mechanisms for the Generation of Visual Complex Cells. *Neuron*, 32(3), 378–380. [https://doi.org/10.1016/S0896-6273\(01\)00497-4](https://doi.org/10.1016/S0896-6273(01)00497-4)
- Cang, J., & Feldheim, D. A. (2013). Developmental Mechanisms of Topographic Map Formation and Alignment [Publisher: Annual Reviews]. *Annual Review of Neuroscience*, 36(Volume 36, 2013), 51–77. <https://doi.org/10.1146/annurev-neuro-062012-170341>
- Cao, A., & Schiller, P. H. (2002). Behavioral assessment of motion parallax and stereopsis as depth cues in rhesus monkeys. *Vision Research*, 42(16), 1953–1961. [https://doi.org/10.1016/S0042-6989\(02\)00117-7](https://doi.org/10.1016/S0042-6989(02)00117-7)
- Cao, A., & Schiller, P. H. (2003). Neural responses to relative speed in the primary visual cortex of rhesus monkey. *Visual Neuroscience*, 20(1), 77–84. <https://doi.org/10.1017/S0952523803201085>
- Carey, D. P., Goodale, M. A., & Sprowl, E. G. (1990). Blindsight in rodents: The use of a ‘high-lever distance cue in gerbils with lesions of primary visual cortex. *Behavioural Brain Research*, 38, 283–289.

- Carter-Dawson, L. D., & Lavail, M. M. (1979). Rods and cones in the mouse retina. I. Structural analysis using light and electron microscopy [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.901880204>]. *Journal of Comparative Neurology*, 188(2), 245–262. <https://doi.org/10.1002/cne.901880204>
- Chalupa, L. M., & Williams, R. W. (Eds.). (2008). *Eye, retina, and visual system of the mouse* [Pages: xii, 754]. Boston Review.
- Chapman, B., Zahs, K. R., & Stryker, M. P. (1991). Relation of cortical cell orientation selectivity to alignment of receptive fields of the geniculocortical afferents that arborize within a single orientation column in ferret visual cortex [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 11(5), 1347–1358. <https://doi.org/10.1523/JNEUROSCI.11-05-01347.1991>
- Chen, A. H., O’Leary, D. J., & Howell, E. R. (2000). Near visual function in young children. Part I: Near point of convergence. Part II: Amplitude of accommodation. Part III: Near heterophoria [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1475-1313.2000.00498.x>]. *Ophthalmic and Physiological Optics*, 20(3), 185–198. <https://doi.org/10.1046/j.1475-1313.2000.00498.x>
- Chen, G., King, J. A., Lu, Y., Cacucci, F., & Burgess, N. (2018). Spatial cell firing during virtual navigation of open arenas by head-restrained mice (L. Colgin, Ed.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, 7, e34789. <https://doi.org/10.7554/eLife.34789>
- Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity [Publisher: Nature Publishing Group]. *Nature*, 499(7458), 295–300. <https://doi.org/10.1038/nature12354>
- Cheney, C., & Crow, R. (1969). Lesion effects on visual cliff performance in the rat. *Psychonomic Science*, 17(3), 165–166. <https://doi.org/10.3758/BF03336499>
- Cheng, X., Qiu, S., Zou, Z., Pu, J., & Xue, X. (2023, June). Understanding Depth Map Progressively: Adaptive Distance Interval Separation for Monocular 3d Object Detection [arXiv:2306.10921 [cs]]. <https://doi.org/10.48550/arXiv.2306.10921>
- Chino, Y., Smith, E., Yoshida, K., Cheng, H., & Hamamoto, J. (1994). Binocular interactions in striate cortical neurons of cats reared with discordant visual inputs. *The Journal of Neuroscience*, 14(8), 5050–5067. <https://doi.org/10.1523/JNEUROSCI.14-08-05050.1994>
- Choi, V., & Priebe, N. J. (2020). Interocular velocity cues elicit vergence eye movements in mice [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 124(2), 623–633. <https://doi.org/10.1152/jn.00697.2019>
- Christensen, A. J., & Pillow, J. W. (2022). Reduced neural activity but improved coding in rodent higher-order visual cortex during locomotion [Publisher: Nature Publishing

- Group]. *Nature Communications*, 13(1), 1676. <https://doi.org/10.1038/s41467-022-29200-z>
- Clark, A. (2016, January). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. OUP USA.
- Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., & Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli. *Perception*, 24(1), 33–47. <https://doi.org/10.1068/p240033>
- Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anticorrelated stereograms: Effects of spatial density and interocular delay. *Vision Research*, 33(14), 1959–1975. [https://doi.org/10.1016/0042-6989\(93\)90021-n](https://doi.org/10.1016/0042-6989(93)90021-n)
- Cornwell, P., Overman, W., Levitsky, C., Shipley, J., & Lezynski, B. (1976). Performance on the visual cliff by cats with marginal gyrus lesions [Place: US Publisher: American Psychological Association]. *Journal of Comparative and Physiological Psychology*, 90(10), 996–1010. <https://doi.org/10.1037/h0077281>
- Cossell, L., Iacaruso, M. F., Muir, D. R., Houlton, R., Sader, E. N., Ko, H., Hofer, S. B., & Mrsic-Flogel, T. D. (2015). Functional organization of excitatory synaptic strength in primary visual cortex [Publisher: Nature Publishing Group]. *Nature*, 518(7539), 399–403. <https://doi.org/10.1038/nature14182>
- Crabtree, J. W., & Riesen, A. H. (1979). Effects of the duration of dark rearing on visually guided behavior in the kitten. *Developmental Psychobiology*, 12(4), 291–303. <https://doi.org/10.1002/dev.420120404>
- Crannell, C. W., & Peters, G. (1970). Monocular and Binocular Estimations of Distance When Knowledge of the Relevant Space is Absent [Publisher: Informa UK Limited]. *The Journal of Psychology*, 76(2), 157–167. <https://doi.org/10.1080/00223980.1970.9916834>
- Cullen, K. E. (2019). Vestibular processing during natural self-motion: Implications for perception and action [Publisher: Nature Publishing Group]. *Nature Reviews Neuroscience*, 20(6), 346–363. <https://doi.org/10.1038/s41583-019-0153-1>
- Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception [Publisher: Nature Publishing Group]. *Nature*, 389(6648), 280–283. <https://doi.org/10.1038/38487>
- Cumming, B. G., & DeAngelis, G. C. (2001). The physiology of stereopsis [Publisher: Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA]. *Annual review of neuroscience*, 24(1), 203–238.
- Dadgarlat, M. C., & Stryker, M. P. (2017). Locomotion Enhances Neural Encoding of Visual Stimuli in Mouse V1. *The Journal of Neuroscience*, 37(14), 3764–3775. <https://doi.org/10.1523/JNEUROSCI.2728-16.2017>
- Dautan, D., Kovács, A., Bayasgalan, T., Diaz-Acevedo, M. A., Pal, B., & Mena-Segovia, J. (2021). Modulation of motor behavior by the mesencephalic locomotor region. *Cell Reports*, 36(8), 109594. <https://doi.org/10.1016/j.celrep.2021.109594>

- David, C. T. (1982). Compensation for height in the control of groundspeed by *Drosophila* in a new, barber's pole wind tunnel. *Journal of comparative physiology*, 147(4), 485–493. <https://doi.org/10.1007/BF00612014>
- Daw, N. (2012, January). Depth Perception. In N. Daw (Ed.), *How Vision Works: The Physiological Mechanisms Behind What We See* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199751617.003.0006>
- De, A., & Horwitz, G. D. (2022). Coding of chromatic spatial contrast by macaque V1 neurons (T. Pasternak, T. Moore, B. R. Conway, & D. Brainard, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, 11, e68133. <https://doi.org/10.7554/eLife.68133>
- de la Cera, E. G., Rodríguez, G., Llorente, L., Schaeffel, F., & Marcos, S. (2006). Optical aberrations in the mouse eye. *Vision Research*, 46(16), 2546–2553. <https://doi.org/10.1016/j.visres.2006.01.011>
- Dean, M. D., & Nachman, M. W. (2009). FASTER FERTILIZATION RATE IN CONSPECIFIC VERSUS HETEROSPECIFIC MATINGS IN HOUSE MICE. *Evolution*, 63(1), 20–28. <https://doi.org/10.1111/j.1558-5646.2008.00499.x>
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993). 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. <https://doi.org/10.1152/jn.1993.69.4.1091>
- DeAngelis, G. C., & Newsome, W. T. (1999). Organization of Disparity-Selective Neurons in Macaque Area MT. *The Journal of Neuroscience*, 19(4), 1398–1415. <https://doi.org/10.1523/jneurosci.19-04-01398.1999>
- DeAngelis, G. C., & Uka, T. (2003). Coding of Horizontal Disparity and Velocity by MT Neurons in the Alert Macaque [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 89(2), 1094–1111. <https://doi.org/10.1152/jn.00717.2002>
- Descartes, R. (1664). *Traite de l'homme* (Vol. 11). Gallimard Paris.
- Ding, J., & Levi, D. M. (2021). A unified model for binocular fusion and depth perception. *Vision Research*, 180, 11–36. <https://doi.org/10.1016/j.visres.2020.11.009>
- Ding, L., & Sharma, G. (2017). Fusing structure from motion and lidar for dense accurate depth map estimation [ISSN: 2379-190X]. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1283–1287. <https://doi.org/10.1109/ICASSP.2017.7952363>
- Dipoppa, M., Ranson, A., Krumin, M., Pachitariu, M., Carandini, M., & Harris, K. D. (2018). Vision and locomotion shape the interactions between neuron types in mouse visual cortex. *Neuron*, 98(3), 602–615.e8. <https://doi.org/10.1016/j.neuron.2018.03.037>
- Domini, F., Caudek, C., & Tassinari, H. (2006). Stereo and motion information are not independently processed by the visual system. *Vision Research*, 46(11), 1707–1723. <https://doi.org/10.1016/j.visres.2005.11.018>

- Dove, H. W. (1841). Ueber die Combination der Eindrücke beider Ohren und beider Augen zu einem Eindruck. *Monatsberichte der Berliner preussischen Akademie der Wissenschaften*, 41, 251–252.
- Dow, B. M. (2002). Orientation and Color Columns in Monkey Visual Cortex. *Cerebral Cortex*, 12(10), 1005–1015. <https://doi.org/10.1093/cercor/12.10.1005>
- Dräger, U. C., & Hubel, D. (1976). Topography of visual and somatosensory projections to mouse superior colliculus [Publisher: American Physiological Society Bethesda, MD]. *Journal of Neurophysiology*, 39(1), 91–101.
- Dräger, U. C., & Olsen, J. F. (1981). Ganglion cell distribution in the retina of the mouse. *Investigative Ophthalmology Visual Science*, 20(3), 285–293.
- Dräger, U. C. (1975). Receptive fields of single cells and topography in mouse visual cortex. *Journal of Comparative Neurology*, 160(3), 269–289. <https://doi.org/10.1002/cne.901600302>
- Dräger, U. C., & Olsen, J. F. (1980). Origins of crossed and uncrossed retinal projections in pigmented and albino mice. *Journal of Comparative Neurology*, 191(3), 383–412. <https://doi.org/10.1002/cne.901910306>
- DSouza, R. D., Wang, Q., Ji, W., Meier, A. M., Kennedy, H., Knoblauch, K., & Burkhalter, A. (2022). Hierarchical and nonhierarchical features of the mouse visual cortical network [Publisher: Nature Publishing Group]. *Nature Communications*, 13(1), 503. <https://doi.org/10.1038/s41467-022-28035-y>
- Du, R., Turner, E., Dzitsiuk, M., Prasso, L., Duarte, I., Dourgarian, J., Afonso, J., Pascoal, J., Gladstone, J., Cruces, N., Izadi, S., Kowdle, A., Tsotsos, K., & Kim, D. (2020). DepthLab: Real-time 3D Interaction with Depth Maps for Mobile Augmented Reality. *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 829–843. <https://doi.org/10.1145/3379337.3415881>
- Durand, J.-B., Trotter, Y., & Celebrini, S. (2010). Privileged Processing of the Straight-Ahead Direction in Primate Area V1 [Publisher: Elsevier]. *Neuron*, 66(1), 126–137. <https://doi.org/10.1016/j.neuron.2010.03.014>
- Ecker, J. L., Dumitrescu, O. N., Wong, K. Y., Alam, N. M., Chen, S.-K., LeGates, T., Renna, J. M., Prusky, G. T., Berson, D. M., & Hattar, S. (2010). Melanopsin-Expressing Retinal Ganglion-Cell Photoreceptors: Cellular Diversity and Role in Pattern Vision. *Neuron*, 67(1), 49–60. <https://doi.org/10.1016/j.neuron.2010.05.023>
- Ellard, C. G., Goodale, M. A., Scorfield, D. M., & Lawrence, C. (1986). Visual cortical lesions abolish the use of motion parallax in the Mongolian gerbil. *Experimental Brain Research*, 64(3), 599–602. <https://doi.org/10.1007/BF00340498>
- Ellard, C. G., Goodale, M. A., & Timney, B. (1984). Distance estimation in the mongolian gerbil: The role of dynamic depth cues. *Behavioural Brain Research*, 14(1), 29–39. [https://doi.org/10.1016/0166-4328\(84\)90017-2](https://doi.org/10.1016/0166-4328(84)90017-2)

- Ellis, E. M., Gauvain, G., Sivyer, B., & Murphy, G. J. (2016). Shared and distinct retinal input to the mouse superior colliculus and dorsal lateral geniculate nucleus [Publisher: American Physiological Society Bethesda, MD]. *Journal of neurophysiology*, 116(2), 602–610.
- Erisken, S., Vaiceliunaite, A., Jurjut, O., Fiorini, M., Katzner, S., & Busse, L. (2014). Effects of Locomotion Extend throughout the Mouse Early Visual System. *Current Biology*, 24(24), 2899–2907. <https://doi.org/https://doi.org/10.1016/j.cub.2014.10.045>
- Evangelista, C., Kraft, P., Dacke, M., Reinhard, J., & Srinivasan, M. V. (2010). The moment before touchdown: Landing manoeuvres of the honeybee *Apis mellifera*. *Journal of Experimental Biology*, 213(2), 262–270. <https://doi.org/10.1242/jeb.037465>
- Fantz, R. L. (1965). Ontogeny of perception. *Behavior of nonhuman primates*, 2, 365–403.
- Farber, J. M., & McConkie, A. B. (1979). Optical motions as information for unsigned depth. *Journal of Experimental Psychology. Human Perception and Performance*, 5(3), 494–500.
- Felleman, D. J., & Van Essen, D. C. (1987). Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 57(4), 889–920. <https://doi.org/10.1152/jn.1987.57.4.889>
- Ferster, D. (1981). A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual cortex [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1981.sp013608>]. *The Journal of Physiology*, 311(1), 623–655. <https://doi.org/10.1113/jphysiol.1981.sp013608>
- Fischer, B., & Krüger, J. (1979). Disparity tuning and binocularity of single neurons in cat visual cortex. *Experimental Brain Research*, 35(1), 1–8. <https://doi.org/10.1007/BF00236780>
- Fisher, S. K., & Ciuffreda, K. J. (1988). Accommodation and Apparent Distance. *Perception*, 17. <https://doi.org/10.1068/p170609>
- Foster, K. H., Gaska, J. P., Nagler, M., & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *The Journal of Physiology*, 365, 331–363. Retrieved December 26, 2024, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1193006/>
- Fox, M. W. (1965). The visual cliff test for the study of visual depth perception in the mouse. *Animal behaviour*, 13(2-3), 232–IN3.
- Fox, R., Aslin, R. N., Shea, S. L., & Dumais, S. T. (1980). Stereopsis in human infants [Publisher: American Association for the Advancement of Science]. *Science*, 207(4428), 323–324.
- Frazor, R. A., Albrecht, D. G., Geisler, W. S., & Crane, A. M. (2004). Visual Cortex Neurons of Monkeys and Cats: Temporal Dynamics of the Spatial Frequency Response

- Function. *Journal of Neurophysiology*, 91(6), 2607–2627. <https://doi.org/10.1152/jn.00858.2003>
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815–836. <https://doi.org/10.1098/rstb.2005.1622>
- Froudarakis, E., Fahey, P. G., Reimer, J., Smirnakis, S. M., Tehovnik, E. J., & Tolias, A. S. (2019). The Visual Cortex in Context [Publisher: Annual Reviews]. *Annual Review of Vision Science*, 5(Volume 5, 2019), 317–339. <https://doi.org/10.1146/annurev-vision-091517-034407>
- Fu, C., Mertz, C., & Dolan, J. M. (2019). LIDAR and Monocular Camera Fusion: On-road Depth Completion for Autonomous Driving. *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 273–278. <https://doi.org/10.1109/ITSC.2019.8917201>
- Fu, J., Pierzchlewicz, P. A., Willeke, K. F., Bashiri, M., Muhammad, T., Diamantaki, M., Froudarakis, E., Restivo, K., Ponder, K., Denfield, G. H., Sinz, F., Tolias, A. S., & Franke, K. (2024). Heterogeneous orientation tuning in the primary visual cortex of mice diverges from Gabor-like receptive fields in primates. *Cell Reports*, 43(8), 114639. <https://doi.org/https://doi.org/10.1016/j.celrep.2024.114639>
- Fu, Y., & Yau, K.-W. (2007). Phototransduction in mouse rods and cones. *Pflügers Archiv : European journal of physiology*, 454(5), 805–819. <https://doi.org/10.1007/s00424-006-0194-y>
- Fu, Y., Tucciarone, J. M., Espinosa, J. S., Sheng, N., Darcy, D. P., Nicoll, R. A., Huang, Z. J., & Stryker, M. P. (2014). A Cortical Circuit for Gain Control by Behavioral State [Publisher: Elsevier]. *Cell*, 156(6), 1139–1152. <https://doi.org/10.1016/j.cell.2014.01.050>
- Gale, S. D., & Murphy, G. J. (2014). Distinct Representation and Distribution of Visual Information by Specific Cell Types in Mouse Superficial Superior Colliculus. *The Journal of Neuroscience*, 34(40), 13458–13471. <https://doi.org/10.1523/JNEUROSCI.2768-14.2014>
- Galletti, C., Battaglini, P. P., & Fattori, P. (1995). Eye Position Influence on the Parieto-occipital Area PO (V6) of the Macaque Monkey [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-9568.1995.tb01047.x>]. *European Journal of Neuroscience*, 7(12), 2486–2501. <https://doi.org/10.1111/j.1460-9568.1995.tb01047.x>
- Galletti, C., & Battaglini, P. (1989). Gaze-dependent visual neurons in area V3A of monkey prestriate cortex. *The Journal of Neuroscience*, 9(4), 1112–1125. <https://doi.org/10.1523/JNEUROSCI.09-04-01112.1989>
- Galletti, C., Squatrito, S., Battaglini, P. P., & Maioli, M. G. (1984). Real-motion cells in the primary visual cortex of macaque monkeys [Publisher: Elsevier]. *Brain research*, 301(1), 95–110.

- Gao, E., DeAngelis, G. C., & Burkhalter, A. (2010). Parallel Input Channels to Mouse Primary Visual Cortex [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 30(17), 5912–5926. <https://doi.org/10.1523/JNEUROSCI.6456-09.2010>
- Garland, T., Gleeson, T. T., Aronovitz, B. A., Richardson, C. S., & Dorm, M. R. (1995). Maximal sprint speeds and muscle fiber composition of wild and laboratory house mice. *Physiology & Behavior*, 58(5), 869–876. [https://doi.org/10.1016/0031-9384\(95\)00148-C](https://doi.org/10.1016/0031-9384(95)00148-C)
- Garrett, M. E., Nauhaus, I., Marshel, J. H., & Callaway, E. M. (2014). Topography and Areal Organization of Mouse Visual Cortex. *The Journal of Neuroscience*, 34(37), 12587. <https://doi.org/10.1523/JNEUROSCI.1124-14.2014>
- Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F., & Marín-Jiménez, M. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6), 2280–2292. <https://doi.org/10.1016/j.patcog.2014.01.005>
- Geng, Y., Schery, L. A., Sharma, R., Dubra, A., Ahmad, K., Libby, R. T., & Williams, D. R. (2011). Optical properties of the mouse eye [Publisher: Optica Publishing Group]. *Biomedical Optics Express*, 2(4), 717–738. <https://doi.org/10.1364/BOE.2.000717>
- Gianfranceschi, L., Siciliano, R., Walls, J., Morales, B., Kirkwood, A., Huang, Z. J., Tonegawa, S., & Maffei, L. (2003). Visual cortex is rescued from the effects of dark rearing by overexpression of BDNF [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, 100(21), 12486–12491. <https://doi.org/10.1073/pnas.1934836100>
- Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, 58(1), 40–51. <https://doi.org/10.1037/h0043883>
- Gibson, E. J., & Walk, R. D. (1960). The Visual Cliff. *Scientific American*, 202(4), 64–71. <https://doi.org/10.1038/scientificamerican0460-64>
- Gibson, J. J., & Carel, W. (1952). Does motion perspective independently produce the impression of a receding surface? [Place: US Publisher: American Psychological Association]. *Journal of Experimental Psychology*, 44(1), 16–18. <https://doi.org/10.1037/h0056030>
- Gilbert, C. D., & Wiesel, T. N. (1983, January). Functional Organization of the Visual Cortex. In J. -. Changeux, J. Glowinski, M. Imbert, & F. E. Bloom (Eds.), *Progress in Brain Research* (pp. 209–218, Vol. 58). Elsevier. [https://doi.org/10.1016/S0079-6123\(08\)60022-9](https://doi.org/10.1016/S0079-6123(08)60022-9)
- Gimeno, J., Morillo, P., Orduña, J. M., & Fernández, M. (2013). A new AR authoring tool using depth maps for industrial procedures. *Computers in Industry*, 64(9), 1263–1271. <https://doi.org/10.1016/j.compind.2013.06.012>

- Glezer, V. D., Tsherbach, T. A., Gauselman, V. E., & Bondarko, V. M. (1982). Spatio-temporal organization of receptive fields of the cat striate cortex. *Biological Cybernetics*, 43(1), 35–49. <https://doi.org/10.1007/BF00337286>
- Glickfeld, L. L., & Olsen, S. R. (2017). Higher-Order Areas of the Mouse Visual Cortex [Publisher: Annu Rev Vis Sci]. *Annual review of vision science*, 3. <https://doi.org/10.1146/annurev-vision-102016-061331>
- Goetz, J., Jessen, Z. F., Jacobi, A., Mani, A., Cooler, S., Greer, D., Kadri, S., Segal, J., Shekhar, K., Sanes, J. R., & Schwartz, G. W. (2022). Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression. *Cell Reports*, 40(2), 111040. <https://doi.org/10.1016/j.celrep.2022.111040>
- Gofman, X., Tocker, G., Weiss, S., Boccara, C. N., Lu, L., Moser, M.-B., Moser, E. I., Morris, G., & Derdikman, D. (2019). Dissociation between Postrhinal Cortex and Downstream Parahippocampal Regions in the Representation of Egocentric Boundaries. *Current Biology*, 29(16), 2751–2757.e4. <https://doi.org/10.1016/j.cub.2019.07.007>
- Gonzalez, F., Krause, F., Perez, R., Alonso, J. M., & Acuna, C. (1993). Binocular matching in monkey visual cortex: Single cell responses to correlated and uncorrelated dynamic random dot stereograms. *Neuroscience*, 52(4), 933–939. [https://doi.org/10.1016/0306-4522\(93\)90540-V](https://doi.org/10.1016/0306-4522(93)90540-V)
- Gonzalez, F., Relova, J. L., Perez, R., Acuña, C., & Alonso, J. M. (1993). Cell responses to vertical and horizontal retinal disparities in the monkey visual cortex. *Neuroscience Letters*, 160(2), 167–170. [https://doi.org/10.1016/0304-3940\(93\)90405-A](https://doi.org/10.1016/0304-3940(93)90405-A)
- Goodale, M. A., Ellard, C. G., & Booth, L. (1990). The role of image size and retinal motion in the computation of absolute distance by the Mongolian gerbil (*Meriones unguiculatus*). *Vision Research*, 30(3), 399–413. [https://doi.org/10.1016/0042-6989\(90\)90082-V](https://doi.org/10.1016/0042-6989(90)90082-V)
- Gordon, J. A., & Stryker, M. P. (1996). Experience-Dependent Plasticity of Binocular Responses in the Primary Visual Cortex of the Mouse [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 16(10), 3274–3286. <https://doi.org/10.1523/JNEUROSCI.16-10-03274.1996>
- Gorski, J. A., Talley, T., Qiu, M., Puelles, L., Rubenstein, J. L. R., & Jones, K. R. (2002). Cortical Excitatory Neurons and Glia, But Not GABAergic Neurons, Are Produced in the Emx1-Expressing Lineage [Publisher: Society for Neuroscience Section: BRIEF COMMUNICATION]. *Journal of Neuroscience*, 22(15), 6309–6314. <https://doi.org/10.1523/JNEUROSCI.22-15-06309.2002>
- Grant, V. W. (1942). Accommodation and convergence in visual space perception [Place: US Publisher: American Psychological Association]. *Journal of Experimental Psychology*, 31(2), 89–104. <https://doi.org/10.1037/h0062359>

- Groschner, L. N., Malis, J. G., Zuidinga, B., & Borst, A. (2022). A biophysical account of multiplication by a single neuron [Publisher: Nature Publishing Group]. *Nature*, 603(7899), 119–123. <https://doi.org/10.1038/s41586-022-04428-3>
- Guan, S.-C., Ju, N.-S., Tao, L., Tang, S.-M., & Yu, C. (2021). Functional organization of spatial frequency tuning in macaque V1 revealed with two-photon calcium imaging. *Progress in Neurobiology*, 205, 102120. <https://doi.org/10.1016/j.pneurobio.2021.102120>
- Guillemot, J.-P., Richer, L., Ptito, M., & Lepore, F. (1993, January). Chapter 17 Disparity coding in the cat: A comparison between areas 17-18 and area 19. In T. P. Hicks, S. Molotchnikoff, & T. Ono (Eds.), *Progress in Brain Research* (pp. 179–187, Vol. 95). Elsevier. [https://doi.org/10.1016/S0079-6123\(08\)60368-4](https://doi.org/10.1016/S0079-6123(08)60368-4)
- Guitchounts, G., Lotter, W., Dapello, J., & Cox, D. (2022, February). Stable 3D Head Direction Signals in the Primary Visual Cortex [Pages: 2020.09.04.283762 Section: New Results]. <https://doi.org/10.1101/2020.09.04.283762>
- Guitchounts, G., Masís, J., Wolff, S. B., & Cox, D. (2020). Encoding of 3D Head Orienting Movements in the Primary Visual Cortex. *Neuron*, 108(3), 512–525.e4. <https://doi.org/10.1016/j.neuron.2020.07.014>
- Güler, A. D., Ecker, J. L., Lall, G. S., Haq, S., Altimus, C. M., Liao, H.-W., Barnard, A. R., Cahill, H., Badea, T. C., Zhao, H., Hankins, M. W., Berson, D. M., Lucas, R. J., Yau, K.-W., & Hattar, S. (2008). Melanopsin cells are the principal conduits for rodcone input to non-image-forming vision [Publisher: Nature Publishing Group]. *Nature*, 453(7191), 102–105. <https://doi.org/10.1038/nature06829>
- Guo, K., & Li, C.-Y. (1997). Eye position-dependent activation of neurones in striate cortex of macaque. *NeuroReport*, 8(6), 1405. Retrieved December 20, 2024, from https://journals.lww.com/neuroreport/abstract/1997/04140/eye_position_dependent_activation_of_neurones_in.17.aspx
- Gur, M., & Snodderly, D. M. (2007). Direction selectivity in V1 of alert monkeys: Evidence for parallel pathways for motion processing. *The Journal of Physiology*, 585(Pt 2), 383–400. <https://doi.org/10.1113/jphysiol.2007.143040>
- Haag, J., Arenz, A., Serbe, E., Gabbiani, F., & Borst, A. (2016). Complementary mechanisms create direction selectivity in the fly (F. Rieke, Ed.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, 5, e17421. <https://doi.org/10.7554/eLife.17421>
- Han, X., Vermaercke, B., & Bonin, V. (2022). Diversity of spatiotemporal coding reveals specialized visual processing streams in the mouse cortex. *Nature Communications*, 13, 3249. <https://doi.org/10.1038/s41467-022-29656-z>
- Han, Y., Kebschull, J. M., Campbell, R. A. A., Cowan, D., Imhof, F., Zador, A. M., & Mrsic-Flogel, T. D. (2018). The logic of single-cell projections from visual cortex [Publisher: Nature Publishing Group]. *Nature*, 556(7699), 51–56. <https://doi.org/10.1038/nature26159>

- Harris, K. D., & Mrsic-Flogel, T. D. (2013). Cortical connectivity and sensory coding [Publisher: Nature Publishing Group]. *Nature*, *503*(7474), 51–58. <https://doi.org/10.1038/nature12654>
- Harris, K. D., & Shepherd, G. M. G. (2015). The neocortical circuit: Themes and variations [Publisher: Nature Publishing Group]. *Nature Neuroscience*, *18*(2), 170–181. <https://doi.org/10.1038/nn.3917>
- Harris, K. D., & Thiele, A. (2011). Cortical state and attention [Publisher: Nature Publishing Group]. *Nature Reviews Neuroscience*, *12*(9), 509–523. <https://doi.org/10.1038/nrn3084>
- Hartline, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina [Publisher: American Physiological Society]. *American Journal of Physiology-Legacy Content*, *121*(2), 400–415. <https://doi.org/10.1152/ajplegacy.1938.121.2.400>
- Hartline, H. K. (1940). The receptive fields of optic nerve fibers [Publisher: American Physiological Society]. *American Journal of Physiology-Legacy Content*, *130*(4), 690–699. <https://doi.org/10.1152/ajplegacy.1940.130.4.690>
- Harwerth, R. S., Smith, E. L., & Crawford, M. L. J. (1996). Motor and sensory fusion in monkeys: Psychophysical measurements [Publisher: Springer Science and Business Media LLC]. *Eye*, *10*(2), 209–216. <https://doi.org/10.1038/eye.1996.48>
- Hatori, M., Le, H., Vollmers, C., Keding, S. R., Tanaka, N., Schmedt, C., Jegla, T., & Panda, S. (2008). Inducible Ablation of Melanopsin-Expressing Retinal Ganglion Cells Reveals Their Central Role in Non-Image Forming Visual Responses [Publisher: Public Library of Science]. *PLOS ONE*, *3*(6), e2451. <https://doi.org/10.1371/journal.pone.0002451>
- Hattar, S., Liao, H.-W., Takao, M., Berson, D. M., & Yau, K.-W. (2002). Melanopsin-Containing Retinal Ganglion Cells: Architecture, Projections, and Intrinsic Photosensitivity [Publisher: American Association for the Advancement of Science]. *Science*, *295*(5557), 1065–1070. <https://doi.org/10.1126/science.1069609>
- Hayashibe, K. (1991). Reversals of visual depth caused by motion parallax. *Perception*, *20*(1), 17–28. <https://doi.org/10.1068/p200017>
- He, Y., Nieto, A. C., Blot, A., & Znamenskiy, P. (2024). A depth map of visual space in the primary visual cortex. <https://doi.org/10.1101/2024.09.27.615442>
- Heesy, C. P. (2004). On the relationship between orbit orientation and binocular visual field overlap in mammals. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology: An Official Publication of the American Association of Anatomists*, *281*(1), 1104–1110.
- Heider, B. (2000). Visual form agnosia: Neural mechanisms and anatomical foundations. *Neurocase*, *6*(1), 1–12. <https://doi.org/10.1080/13554790008402753>
- Hein, A., Held, R., & Gower, E. C. (1970). Development and segmentation of visually controlled movement by selective exposure during rearing [Place: US Publisher:

- American Psychological Association]. *Journal of Comparative and Physiological Psychology*, 73(2), 181–187. <https://doi.org/10.1037/h0030247>
- Held, R., & Hein, A. (1963). Movement-produced stimulation in the development of visually guided behavior [Place: US Publisher: American Psychological Association]. *Journal of Comparative and Physiological Psychology*, 56(5), 872–876. <https://doi.org/10.1037/h0040546>
- Helmholtz, H. v. (1868). *Die neueren Fortschritte in der Theorie des Sehens* [Google-Books-ID: E9aOtwAACAAJ].
- Helmholtz, H. v. (1925). *Helmholtz's Treatise on Physiological Optics* [Google-Books-ID: 1EzPAAAAMAAJ]. Optical Society of America.
- Herrera, E., Brown, L., Aruga, J., Rachel, R. A., Dolen, G., Mikoshiba, K., Brown, S., & Mason, C. A. (2003). Zic2 Patterns Binocular Vision by Specifying the Uncrossed Retinal Projection [Publisher: Elsevier]. *Cell*, 114(5), 545–557. [https://doi.org/10.1016/S0092-8674\(03\)00684-6](https://doi.org/10.1016/S0092-8674(03)00684-6)
- Hillebrand, F. (1894). Das Verhältniss von accommodation und Konvergenz zur Tiefenlokalisation. *Zeitschrift für Psychologie*, 7, 97–151.
- Hinkle, D. A., & Connor, C. E. (2002). Three-dimensional orientation tuning in macaque area V4 [Publisher: Nature Publishing Group]. *Nature Neuroscience*, 5(7), 665–670. <https://doi.org/10.1038/nn875>
- Hinman, J. R., Chapman, G. W., & Hasselmo, M. E. (2019). Neuronal representation of environmental boundaries in egocentric coordinates [Publisher: Nature Publishing Group]. *Nature Communications*, 10(1), 2772. <https://doi.org/10.1038/s41467-019-10722-y>
- Hirsch, M. J., & Weymouth, F. W. (1948). Distance discrimination; effect on threshold of lateral separation of the test objects [Publisher: Arch Ophthal]. *Archives of ophthalmology (Chicago, Ill. : 1929)*, 39(2). Retrieved July 26, 2025, from <https://pubmed.ncbi.nlm.nih.gov/18868828/>
- Holmgren, C. D., Stahr, P., Wallace, D. J., Voit, K.-M., Matheson, E. J., Sawinski, J., Bassetto, G., & Kerr, J. N. (2021). Visual pursuit behavior in mice maintains the pursued prey on the retinal region with least optic flow. *eLife*, 10, e70838. <https://doi.org/10.7554/eLife.70838>
- Horridge, G. A. (1986). A Theory of Insect Vision: Velocity Parallax [Publisher: Royal Society]. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, 229(1254), 13–27. Retrieved November 21, 2024, from <https://www.jstor.org/stable/36074>
- Howard, C. H. J. (1919). A Test for the Judgment of Distance [Publisher: Elsevier BV]. *American Journal of Ophthalmology*, 2(9), 656–675. [https://doi.org/10.1016/s0002-9394\(19\)90180-2](https://doi.org/10.1016/s0002-9394(19)90180-2)
- Howard, I. P. (2012a, February). Depth from accommodation and vergence. In I. P. Howard (Ed.), *Perceiving in Depth: Volume 3 Other Mechanisms of Depth Per-*

- ception* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199764167.003.0007>
- Howard, I. P. (2012b, February). Depth from motion parallax. In I. P. Howard (Ed.), *Perceiving in Depth: Volume 3 Other Mechanisms of Depth Perception* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199764167.003.0116>
- Howard, I. P. (2012c, February). Visual Depth Perception in the Animal Kingdom. In I. P. Howard (Ed.), *Perceiving in Depth: Volume 3 Other Mechanisms of Depth Perception* (p. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199764167.003.0231>
- Howard, I. P., & Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford University Press, USA.
- Howard, I., & Rogers, B. (2002). *Seeing in depth. Volume 2: Depth perception*. I. Porteus, Toronto, Canada.
- Hoy, J. L., Bishop, H. I., & Niell, C. M. (2019). Defined Cell Types in Superior Colliculus Make Distinct Contributions to Prey Capture Behavior in the Mouse [Publisher: Elsevier]. *Current Biology*, 29(23), 4130–4138.e5. <https://doi.org/10.1016/j.cub.2019.10.017>
- Hoy, J. L., & Niell, C. M. (2015). Layer-Specific Refinement of Visual Cortex Function after Eye Opening in the Awake Mouse [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 35(8), 3370–3383. <https://doi.org/10.1523/JNEUROSCI.3174-14.2015>
- Høydal, Ø. A., Skytøen, E. R., Andersson, S. O., Moser, M.-B., & Moser, E. I. (2019). Object-vector coding in the medial entorhinal cortex [Publisher: Nature Publishing Group]. *Nature*, 568(7752), 400–404. <https://doi.org/10.1038/s41586-019-1077-7>
- Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3), 574–591. <https://doi.org/10.1113/jphysiol.1959.sp006308>
- Hubel, D. H., & Wiesel, T. N. (1961). Integrative action in the cat's lateral geniculate body. *The Journal of Physiology*, 155(2), 385–398. <https://doi.org/10.1113/jphysiol.1961.sp006635>
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1962.sp006837>]. *The Journal of Physiology*, 160(1), 106–154. <https://doi.org/10.1113/jphysiol.1962.sp006837>
- Hubel, D. H., & Wiesel, T. N. (1963). Shape and arrangement of columns in cat's striate cortex. *The Journal of Physiology*, 165(3), 559–568. <https://doi.org/10.1113/jphysiol.1963.sp007079>

- Hubel, D. H., & Wiesel, T. N. (1969). Anatomical Demonstration of Columns in the Monkey Striate Cortex [Publisher: Nature Publishing Group]. *Nature*, *221*(5182), 747–750. <https://doi.org/10.1038/221747a0>
- Hubel, D. H., & Wiesel, T. N. (1965). BINOCULAR INTERACTION IN STRIATE CORTEX OF KITTENS REARED WITH ARTIFICIAL SQUINT. *Journal of Neurophysiology*, *28*(6), 1041–1059. <https://doi.org/10.1152/jn.1965.28.6.1041>
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex [Publisher: Wiley Online Library]. *The Journal of physiology*, *195*(1), 215–243.
- Hubel, D. H., & Wiesel, T. N. (1970). Stereoscopic vision in macaque monkey: Cells sensitive to binocular depth in area 18 of the macaque monkey cortex [Publisher: Nature Publishing Group UK London]. *Nature*, *225*(5227), 41–42.
- Hubel, D. H., & Wiesel, T. N. (1972). Laminar and columnar distribution of geniculocortical fibers in the macaque monkey [Publisher: Wiley Online Library]. *Journal of Comparative Neurology*, *146*(4), 421–450.
- Hubel, D. H., & Wiesel, T. N. (1974). Sequence regularity and geometry of orientation columns in the monkey striate cortex [Publisher: Wiley Online Library]. *Journal of Comparative Neurology*, *158*(3), 267–293.
- Hubel, D. H., Wiesel, T. N., & Stryker, M. P. (1977). Orientation columns in macaque monkey visual cortex demonstrated by the 2-deoxyglucose autoradiographic technique [Publisher: Nature Publishing Group]. *Nature*, *269*(5626), 328–330. <https://doi.org/10.1038/269328a0>
- Hubel, D. H., Wiesel, T. N., Yeagle, E. M., Lafer-Sousa, R., & Conway, B. R. (2015). Binocular Stereoscopy in Visual Areas V-2, V-3, and V-3A of the Macaque Monkey. *Cerebral Cortex*, *25*(4), 959–971. <https://doi.org/10.1093/cercor/bht288>
- Hubel, D., & Livingstone, M. (1987). Segregation of form, color, and stereopsis in primate area 18. *The Journal of Neuroscience*, *7*(11), 3378–3415. <https://doi.org/10.1523/JNEUROSCI.07-11-03378.1987>
- Ibrahim, M. M., Liu, Q., Khan, R., Yang, J., Adeli, E., & Yang, Y. (2020). Depth map artefacts reduction: A review [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1049/iet-ipr.2019.1622>]. *IET Image Processing*, *14*(12), 2630–2644. <https://doi.org/10.1049/iet-ipr.2019.1622>
- Iriki, A., Tanaka, M., & Iwamura, Y. (1996). Attention-induced neuronal activity in the monkey somatosensory cortex revealed by pupillometrics. *Neuroscience Research*, *25*(2), 173–181. [https://doi.org/10.1016/0168-0102\(96\)01043-7](https://doi.org/10.1016/0168-0102(96)01043-7)
- Issa, N. P., Trepel, C., & Stryker, M. P. (2000). Spatial frequency maps in cat visual cortex [Publisher: Soc Neuroscience]. *Journal of Neuroscience*, *20*(22), 8504–8514.
- Ito, S., & Feldheim, D. A. (2018). The mouse superior colliculus: An emerging model for studying circuit formation and function [Publisher: Frontiers Media SA]. *Frontiers in neural circuits*, *12*, 10.

- Jackson, J., Ayzenshtat, I., Karnani, M. M., & Yuste, R. (2016). VIP+ interneurons control neocortical activity across brain states. *Journal of Neurophysiology*, *115*(6), 3008–3017. <https://doi.org/10.1152/jn.01124.2015>
- Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2003). At Least at the Level of Inferior Temporal Cortex, the Stereo Correspondence Problem Is Solved [Publisher: Elsevier]. *Neuron*, *37*(4), 693–701. [https://doi.org/10.1016/S0896-6273\(03\)00023-0](https://doi.org/10.1016/S0896-6273(03)00023-0)
- Jaubert-Miazza, L., Green, E., Lo, F.-S., Bui, K., Mills, J., & Guido, W. (2005). Structural and functional composition of the developing retinogeniculate pathway in the mouse. *Visual Neuroscience*, *22*(5), 661–676. <https://doi.org/10.1017/S0952523805225154>
- Ji, X.-y., Zingg, B., Mesik, L., Xiao, Z., Zhang, L. I., & Tao, H. W. (2016). Thalamocortical Innervation Pattern in Mouse Auditory and Visual Cortex: Laminar and Cell-Type Specificity. *Cerebral Cortex*, *26*(6), 2612–2625. <https://doi.org/10.1093/cercor/bhv099>
- Jiang, X., Shen, S., Cadwell, C. R., Berens, P., Sinz, F., Ecker, A. S., Patel, S., & Tolias, A. S. (2015). Principles of connectivity among morphologically defined cell types in adult neocortex [Publisher: American Association for the Advancement of Science]. *Science*, *350*(6264), aac9462. <https://doi.org/10.1126/science.aac9462>
- Johnson, K. P., Fitzpatrick, M. J., Zhao, L., Wang, B., McCracken, S., Williams, P. R., & Kerschensteiner, D. (2021). Cell-type-specific binocular vision guides predation in mice. *Neuron*, *109*(9), 1527–1539.e4. <https://doi.org/10.1016/j.neuron.2021.03.010>
- Jordan, C. (2013). Feature Extraction from Depth Maps for Object Recognition. *Technical Report*.
- Jordan, R., & Keller, G. B. (2020). Opposing Influence of Top-down and Bottom-up Input on Excitatory Layer 2/3 Neurons in Mouse Primary Visual Cortex. *Neuron*.
- Juavinett, A. L., Kim, E. J., Collins, H. C., & Callaway, E. M. (2020). A systematic topographical relationship between mouse lateral posterior thalamic neurons and their visual cortical projection targets. *The Journal of comparative neurology*, *528*(1), 95–107. <https://doi.org/10.1002/cne.24737>
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago, University of Chicago Press.
- Kalatsky, V. A., & Stryker, M. P. (2003). New Paradigm for Optical Imaging Temporally Encoded Maps of Intrinsic Signal. *Neuron*, *38*(4), 529–545. [https://doi.org/10.1016/s0896-6273\(03\)00286-1](https://doi.org/10.1016/s0896-6273(03)00286-1)
- Kapustjansky, A., Chittka, L., & Spaethe, J. (2010). Bees use three-dimensional information to improve target detection. *Naturwissenschaften*, *97*(2), 229–233. <https://doi.org/10.1007/s00114-009-0627-5>
- Keller, G. B., Bonhoeffer, T., & Hübener, M. (2012). Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. *Neuron*, *74*(5), 809–815. <https://doi.org/10.1016/j.neuron.2012.03.040>

- Keller, G. B., & Mrsic-Flogel, T. D. (2018). Predictive Processing: A Canonical Cortical Computation [Publisher: Europe PMC Funders]. *Neuron*, 100(2), 424. <https://doi.org/10.1016/j.neuron.2018.10.003>
- Kim, H. R., Angelaki, D. E., & DeAngelis, G. C. (2015a). A Functional Link between MT Neurons and Depth Perception Based on Motion Parallax. *The Journal of Neuroscience*, 35(6), 2766–2777. <https://doi.org/10.1523/jneurosci.3134-14.2015>
- Kim, H. R., Angelaki, D. E., & DeAngelis, G. C. (2015b). A novel role for visual perspective cues in the neural computation of depth [Publisher: Nature Publishing Group]. *Nature Neuroscience*, 18(1), 129–137. <https://doi.org/10.1038/nn.3889>
- Kim, H. R., Angelaki, D. E., & DeAngelis, G. C. (2022). A neural mechanism for detecting object motion during self-motion (D. A. Clark, T. Moore, A. S. Nandy, & O. W. Layton, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, 11, e74971. <https://doi.org/10.7554/eLife.74971>
- Kim, H. R., Angelaki, D. E., & DeAngelis, G. C. (2016). The neural basis of depth perception from motion parallax. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1697), 20150256. <https://doi.org/10.1098/rstb.2015.0256>
- Kim, M.-H., Znamenskiy, P., Iacaruso, M. F., & Mrsic-Flogel, T. D. (2018). Segregated subnetworks of intracortical projection neurons in primary visual cortex. *Neuron*, 100(6), 1313–1321.
- Koehler, C. C., Hall, L. M., Hellmer, C. B., & Ichinose, T. (2019). Using Looming Visual Stimuli to Evaluate Mouse Vision. *Journal of visualized experiments : JoVE*, (148), 10.3791/59766. <https://doi.org/10.3791/59766>
- Körding, K. P., & Wolpert, D. M. (2004). Bayesian integration in sensorimotor learning [Publisher: Nature Publishing Group]. *Nature*, 427(6971), 244–247. <https://doi.org/10.1038/nature02169>
- Kral, K. (2003). Behavioural analytical studies of the role of head movements in depth perception in insects, birds and mammals. *Behavioural Processes*, 64(1), 1–12. [https://doi.org/10.1016/S0376-6357\(03\)00054-8](https://doi.org/10.1016/S0376-6357(03)00054-8)
- Kravitz, D. J., Saleem, K. S., Baker, C. I., & Mishkin, M. (2011). A new neural framework for visuospatial processing [Publisher: Nature Publishing Group]. *Nature Reviews Neuroscience*, 12(4), 217–230. <https://doi.org/10.1038/nrn3008>
- Krug, K., Cumming, B. G., & Parker, A. J. (2004). Comparing Perceptual Signals of Single V5/MT Neurons in Two Binocular Depth Tasks [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 92(3), 1586–1596. <https://doi.org/10.1152/jn.00851.2003>
- Kshirsagar, K. P., Tathod, A., Dudhate, R., Yadav, A., & Tarte, R. (2022). Camera Calibration Using Robust Intrinsic and Extrinsic Parameters [Num Pages: 11]. In *Artificial Intelligence in Information and Communication Technologies, Healthcare and Education*. Chapman; Hall/CRC.

- Kuan, A. T., Bondanelli, G., Driscoll, L. N., Han, J., Kim, M., Hildebrand, D. G. C., Graham, B. J., Wilson, D. E., Thomas, L. A., Panzeri, S., Harvey, C. D., & Lee, W.-C. A. (2024). Synaptic wiring motifs in posterior parietal cortex support decision-making [Publisher: Nature Publishing Group]. *Nature*, 627(8003), 367–373. <https://doi.org/10.1038/s41586-024-07088-7>
- Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16(1), 37–68. <https://doi.org/10.1152/jn.1953.16.1.37>
- La Chioma, A., Bonhoeffer, T., & Hübener, M. (2019). Area-specific mapping of binocular disparity across mouse visual cortex. *Current Biology*, 29(17), 2954–2960.e5. <https://doi.org/10.1016/j.cub.2019.07.037>
- La Chioma, A., Bonhoeffer, T., & Hübener, M. (2020). Disparity Sensitivity and Binocular Integration in Mouse Visual Cortex Areas [Publisher: Society for Neuroscience Section: Research Articles]. *Journal of Neuroscience*, 40(46), 8883–8899. <https://doi.org/10.1523/JNEUROSCI.1060-20.2020>
- Lagae, L., Raiguel, S., & Orban, G. A. (1993). Speed and direction selectivity of macaque middle temporal neurons. *Journal of Neurophysiology*, 69(1), 19–39. <https://doi.org/10.1152/jn.1993.69.1.19>
- Lashley, K. S., & Russell, J. T. (1934). The Mechanism of Vision. XI. A Preliminary Test of Innate Organization. *The Pedagogical Seminary and Journal of Genetic Psychology*, 45(1), 136–144. <https://doi.org/10.1080/08856559.1934.10534252>
- LeDue, E. E., Zou, M. Y., & Crowder, N. A. (2012). Spatiotemporal tuning in mouse primary visual cortex. *Neuroscience Letters*, 528(2), 165–169. <https://doi.org/10.1016/j.neulet.2012.09.006>
- Lee, A. M., Hoy, J. L., Bonci, A., Wilbrecht, L., Stryker, M. P., & Niell, C. M. (2014). Identification of a Brainstem Circuit Regulating Visual Cortical State in Parallel with Locomotion [Publisher: Elsevier]. *Neuron*, 83(2), 455–466. <https://doi.org/10.1016/j.neuron.2014.06.031>
- Lee, S.-H., & Dan, Y. (2012). Neuromodulation of Brain States [Publisher: Elsevier]. *Neuron*, 76(1), 209–222. <https://doi.org/10.1016/j.neuron.2012.09.012>
- Lehrer, M., Srinivasan, M. V., Zhang, S. W., & Horridge, G. A. (1988). Motion cues provide the bee’s visual world with a third dimension [Publisher: Nature Publishing Group]. *Nature*, 332(6162), 356–357. <https://doi.org/10.1038/332356a0>
- Leinweber, M., Ward, D. R., Sobczak, J. M., Attinger, A., & Keller, G. B. (2017). A sensorimotor circuit in mouse cortex for visual flow predictions. *Neuron*, 95(6), 1420–1432.
- LeVay, S., Hubel, D. H., & Wiesel, T. N. (1975). The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.901590408>]. *Journal of Comparative Neurology*, 159(4), 559–575. <https://doi.org/10.1002/cne.901590408>

- Lever, C., Burton, S., Jeewajee, A., O’Keefe, J., & Burgess, N. (2009). Boundary Vector Cells in the Subiculum of the Hippocampal Formation [Publisher: Society for Neuroscience Section: Brief Communications]. *Journal of Neuroscience*, 29(31), 9771–9777. <https://doi.org/10.1523/JNEUROSCI.1319-09.2009>
- Li, P., Farin, D., Klein Gunnewiek, R., & With, d., P.H.N. (2006). On creating depth maps from monoscopic video using structure from motion (R. Lagendijk, J. Weber, & v. d. Berg F.A., Eds.) [Publisher: Werkgemeenschap voor Informatie- en Communicatietheorie (WIC)]. *Proc. 27th Symposium on Information Theory in the Benelux, June 8-9, 2006, Noordwijk, The Netherlands*, 85–91.
- Liao, M., Lu, F., Zhou, D., Zhang, S., Li, W., & Yang, R. (2020). DVI: Depth Guided Video Inpainting for Autonomous Driving. In A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm (Eds.), *Computer Vision ECCV 2020* (pp. 1–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-58589-1_1
- Lien, A. D., & Scanziani, M. (2018). Cortical direction selectivity emerges at convergence of thalamic synapses [Publisher: Nature Publishing Group]. *Nature*, 558(7708), 80–86. <https://doi.org/10.1038/s41586-018-0148-5>
- Liu, B.-h., Li, Y.-t., Ma, W.-p., Pan, C.-j., Zhang, L. I., & Tao, H. W. (2011). Broad Inhibition Sharpens Orientation Selectivity By Expanding Input Dynamic Range in Mouse Simple Cells. *Neuron*, 71(3), 542–554. <https://doi.org/10.1016/j.neuron.2011.06.017>
- Liu, X., Huang, H., Snutch, T. P., Cao, P., Wang, L., & Wang, F. (2022). The Superior Colliculus: Cell Types, Connectivity, and Behavior. *Neuroscience Bulletin*, 38(12), 1519–1540. <https://doi.org/10.1007/s12264-022-00858-1>
- Liu, Y., Vogels, R., & Orban, G. A. (2004). Convergence of Depth from Texture and Depth from Disparity in Macaque Inferior Temporal Cortex. *The Journal of Neuroscience*, 24(15), 3795–3800. <https://doi.org/10.1523/JNEUROSCI.0150-04.2004>
- Liu, Y., Shi, X., Li, Y., & Zhao, K. (2012). The influences of dark rearing on the transmission characteristics of layer II/III pyramidal cells during the critical period. *Brain Research*, 1457, 26–32. <https://doi.org/10.1016/j.brainres.2012.03.062>
- Lopes, G., Bonacchi, N., Frazão, J., Neto, J. P., Atallah, B. V., Soares, S., Moreira, L., Matias, S., Itskov, P. M., Correia, P. A., Medina, R. E., Calcaterra, L., Dreosti, E., Paton, J. J., & Kampff, A. R. (2015). Bonsai: An event-based framework for processing and controlling data streams. *Frontiers in Neuroinformatics*, 9, 7. <https://doi.org/10.3389/fninf.2015.00007>
- Lopes, G., Farrell, K., Horrocks, E. A., Lee, C.-Y., Morimoto, M. M., Muzzu, T., Papanikolaou, A., Rodrigues, F. R., Wheatcroft, T., Zucca, S., Solomon, S. G., & Saleem, A. B. (2021). Creating and controlling visual environments using BonVision (C. I. Baker, J. P. Newman, A. Maia Chagas, & S. A. Koay, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, 10, e65541. <https://doi.org/10.7554/eLife.65541>

- Lu, J., Behbahani, A. H., Hamburg, L., Westeinde, E. A., Dawson, P. M., Lyu, C., Maimon, G., Dickinson, M. H., Druckmann, S., & Wilson, R. I. (2022). Transforming representations of movement from body- to world-centric space [Publisher: Nature Publishing Group]. *Nature*, *601*(7891), 98–104. <https://doi.org/10.1038/s41586-021-04191-x>
- Lucas, R., Hattar, S., Takao, M., Berson, D., Foster, R., & Yau, K.-W. (2003). Diminished pupillary light reflex at high irradiances in melanopsin-knockout mice. *Science*, *299*(5604), 245–247. <https://doi.org/10.1126/science.1077293>
- Luo, L., Callaway, E. M., & Svoboda, K. (2008). Genetic Dissection of Neural Circuits [Publisher: Elsevier]. *Neuron*, *57*(5), 634–660. <https://doi.org/10.1016/j.neuron.2008.01.002>
- Lyu, C., Abbott, L. F., & Maimon, G. (2022). Building an allocentric travelling direction signal via vector computation [Number: 7891 Publisher: Nature Publishing Group]. *Nature*, *601*(7891), 92–97. <https://doi.org/10.1038/s41586-021-04067-0>
- Ma, X., Wang, Z., Li, H., Zhang, P., Fan, X., & Ouyang, W. (2021, March). Accurate Monocular Object Detection via Color-Embedded 3D Reconstruction for Autonomous Driving [arXiv:1903.11444 [cs]]. <https://doi.org/10.48550/arXiv.1903.11444>
- Madisen, L., Garner, A. R., Shimaoka, D., Chuong, A. S., Klapoetke, N. C., Li, L., van der Bourg, A., Niino, Y., Egolf, L., Monetti, C., Gu, H., Mills, M., Cheng, A., Tasic, B., Nguyen, T. N., Sunkin, S. M., Benucci, A., Nagy, A., Miyawaki, A., ... Zeng, H. (2015). Transgenic Mice for Intersectional Targeting of Neural Sensors and Effectors with High Specificity and Performance. *Neuron*, *85*(5), 942–958. <https://doi.org/10.1016/j.neuron.2015.02.022>
- Mante, V., & Carandini, M. (2005). Mapping of Stimulus Energy in Primary Visual Cortex [Publisher: American Physiological Society]. *Journal of Neurophysiology*, *94*(1), 788–798. <https://doi.org/10.1152/jn.01094.2004>
- Marotta, J. J., Behrmann, M., & Goodale, M. A. (1997). The removal of binocular cues disrupts the calibration of grasping in patients with visual form agnosia: *Experimental Brain Research*, *116*(1), 113–121. <https://doi.org/10.1007/PL00005731>
- Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT Press.
- Marshel, J. H., Garrett, M. E., Nauhaus, I., & Callaway, E. M. (2011). Functional specialization of seven mouse visual cortical areas. *Neuron*, *72*(6), 1040–1054.
- Martin, K. A. C. (1994). A Brief History of the Feature Detector. *Cerebral Cortex*, *4*(1), 1–7. <https://doi.org/10.1093/cercor/4.1.1>
- Martin, P. R. (1986). The projection of different retinal ganglion cell classes to the dorsal lateral geniculate nucleus in the hooded rat. *Experimental Brain Research*, *62*(1), 77–88. <https://doi.org/10.1007/BF00237404>

- Martins, D. M., Manda, J. M., Goard, M. J., & Parker, P. R. L. (2024). Building egocentric models of local space from retinal input. *Current Biology*, *34*(23), R1185–R1202. <https://doi.org/10.1016/j.cub.2024.10.057>
- Masland, R. H. (2012). The Neuronal Organization of the Retina. *Neuron*, *76*(2), 266–280. <https://doi.org/10.1016/j.neuron.2012.10.002>
- 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 [Publisher: Nature Publishing Group]. *Nature Neuroscience*, *21*(9), 1281–1289. <https://doi.org/10.1038/s41593-018-0209-y>
- Maunsell, J. H., & Van Essen, D. C. (1983). 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. <https://doi.org/10.1152/jn.1983.49.5.1127>
- Mayford, M., Bach, M. E., Huang, Y.-Y., Wang, L., Hawkins, R. D., & Kandel, E. R. (1996). Control of Memory Formation Through Regulated Expression of a CaMKII Transgene [Publisher: American Association for the Advancement of Science]. *Science*, *274*(5293), 1678–1683. <https://doi.org/10.1126/science.274.5293.1678>
- Mazer, J. A., Vinje, W. E., McDermott, J., Schiller, P. H., & Gallant, J. L. (2002). Spatial frequency and orientation tuning dynamics in area V1 [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *99*(3), 1645–1650. <https://doi.org/10.1073/pnas.022638499>
- McBride, E. G., Lee, S.-Y. J., & Callaway, E. M. (2019). Local and Global Influences of Visual Spatial Selection and Locomotion in Mouse Primary Visual Cortex. *Current biology: CB*, *29*(10), 1592–1605.e5. <https://doi.org/10.1016/j.cub.2019.03.065>
- McKee, S. P., & Taylor, D. G. (2010). The precision of binocular and monocular depth judgments in natural settings. *Journal of vision*, *10*(10), 5. <https://doi.org/10.1167/10.10.5>
- McLaughlin, T., & O’Leary, D. D. M. (2005). MOLECULAR GRADIENTS AND DEVELOPMENT OF RETINOTOPIC MAPS [Publisher: Annual Reviews]. *Annual Review of Neuroscience*, *28*(Volume 28, 2005), 327–355. <https://doi.org/10.1146/annurev.neuro.28.061604.135714>
- Merriam, E. P., Gardner, J. L., Movshon, J. A., & Heeger, D. J. (2013). Modulation of Visual Responses by Gaze Direction in Human Visual Cortex [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, *33*(24), 9879–9889. <https://doi.org/10.1523/JNEUROSCI.0500-12.2013>
- Meyer, A. F., OKeefe, J., & Poort, J. (2020). Two Distinct Types of Eye-Head Coupling in Freely Moving Mice. *Current Biology*, *30*(11), 2116–2130.e6. <https://doi.org/10.1016/j.cub.2020.04.042>
- Meyer, A. F., Poort, J., OKeefe, J., Sahani, M., & Linden, J. F. (2018). A Head-Mounted Camera System Integrates Detailed Behavioral Monitoring with Multichannel Elec-

- trophysiology in Freely Moving Mice [Publisher: Elsevier]. *Neuron*, 100(1), 46–60.e7. <https://doi.org/10.1016/j.neuron.2018.09.020>
- Meyer, P. M., Anderson, R. A., & Braun, M. G. (1966). Visual cliff preferences following lesions of the visual neocortex in cats and rats. *Psychonomic Science*, 4(7), 269–270.
- Meyer, P. M. (1963). Analysis of visual behavior in cats with extensive neocortical ablations. *Journal of Comparative and Physiological Psychology*, 56(2), 397–401. <https://doi.org/https://doi.org/10.1037/h0047822>
- Michaïel, A. M., Abe, E. T., & Niell, C. M. (2020). Dynamics of gaze control during prey capture in freely moving mice (M. Spering, L. L. Colgin, T. Isa, A. Leonardo, & A. K. Churchland, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, 9, e57458. <https://doi.org/10.7554/eLife.57458>
- Miles, F. (1997). Visual stabilization of the eyes in primates. *Current Opinion in Neurobiology*, 7(6), 867–871. [https://doi.org/10.1016/S0959-4388\(97\)80148-1](https://doi.org/10.1016/S0959-4388(97)80148-1)
- Miles, F. (1998). The neural processing of 3-D visual information: Evidence from eye movements [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1460-9568.1998.00112.x>]. *European Journal of Neuroscience*, 10(3), 811–822. <https://doi.org/10.1046/j.1460-9568.1998.00112.x>
- Mineault, P. J., Tring, E., Trachtenberg, J. T., & Ringach, D. L. (2016). Enhanced Spatial Resolution During Locomotion and Heightened Attention in Mouse Primary Visual Cortex [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 36(24), 6382–6392. <https://doi.org/10.1523/JNEUROSCI.0430-16.2016>
- Miura, S. K., & Scanziani, M. (2022). Distinguishing externally from saccade-induced motion in visual cortex. *Nature*, 610(7930), 135–142. <https://doi.org/10.1038/s41586-022-05196-w>
- Mon-Williams, M., Tresilian, J. R., McIntosh, R. D., & Milner, D. A. (2001). Monocular and binocular distance cues: Insights from visual form agnosia I (of III). *Experimental Brain Research*, 139(2), 127–136. <https://doi.org/10.1007/s002210000657>
- Morales, B., Choi, S.-Y., & Kirkwood, A. (2002). Dark Rearing Alters the Development of GABAergic Transmission in Visual Cortex. *The Journal of Neuroscience*, 22(18), 8084–8090. <https://doi.org/10.1523/JNEUROSCI.22-18-08084.2002>
- Morimoto, M. M., Uchishiba, E., & Saleem, A. B. (2021). Organization of feedback projections to mouse primary visual cortex. *iScience*, 24(5), 102450. <https://doi.org/10.1016/j.isci.2021.102450>
- Morin, L. P., & Studholme, K. M. (2014). Retinofugal projections in the mouse [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.23635>]. *Journal of Comparative Neurology*, 522(16), 3733–3753. <https://doi.org/10.1002/cne.23635>

- Morris, A. P., & Krekelberg, B. (2019). A Stable Visual World in Primate Primary Visual Cortex [Publisher: Elsevier]. *Current Biology*, 29(9), 1471–1480.e6. <https://doi.org/10.1016/j.cub.2019.03.069>
- Morris, A. P., Kubischik, M., Hoffmann, K.-P., Krekelberg, B., & Bremmer, F. (2012). Dynamics of Eye-Position Signals in the Dorsal Visual System. *Current Biology*, 22(3), 173–179. <https://doi.org/10.1016/j.cub.2011.12.032>
- Mrsic-Flogel, T. D., Hofer, S. B., Ohki, K., Reid, R. C., Bonhoeffer, T., & Hübener, M. (2007). Homeostatic Regulation of Eye-Specific Responses in Visual Cortex during Ocular Dominance Plasticity [Publisher: Elsevier]. *Neuron*, 54(6), 961–972. <https://doi.org/10.1016/j.neuron.2007.05.028>
- Müller, C. M. (1990). Dark-rearing retards the maturation of astrocytes in restricted layers of cat visual cortex. *Glia*, 3(6), 487–494. <https://doi.org/10.1002/glia.440030607>
- Müller, K., Merkle, P., & Wiegand, T. (2011). 3-D Video Representation Using Depth Maps [Conference Name: Proceedings of the IEEE]. *Proceedings of the IEEE*, 99(4), 643–656. <https://doi.org/10.1109/JPROC.2010.2091090>
- Murgas, K. A., Wilson, A. M., Michael, V., & Glickfeld, L. L. (2019, May). Unique spatial integration in mouse primary visual cortex and higher visual areas [Pages: 643007 Section: New Results]. <https://doi.org/10.1101/643007>
- Muzzu, T., & Saleem, A. B. (2021). Feature selectivity can explain mismatch signals in mouse visual cortex [Publisher: Cell Rep]. *Cell reports*, 37(1). <https://doi.org/10.1016/j.celrep.2021.109772>
- Nadler, J. W., Angelaki, D. E., & DeAngelis, G. C. (2008). A neural representation of depth from motion parallax in macaque visual cortex. *Nature*, 452(7187), 642–645.
- Nadler, J. W., Barbash, D., Kim, H. R., Shimpi, S., Angelaki, D. E., & DeAngelis, G. C. (2013). Joint Representation of Depth from Motion Parallax and Binocular Disparity Cues in Macaque Area MT. *The Journal of Neuroscience*, 33(35), 14061–14074. <https://doi.org/10.1523/JNEUROSCI.0251-13.2013>
- Nadler, J. W., Nawrot, M., Angelaki, D. E., & DeAngelis, G. C. (2009). MT neurons combine visual motion with a smooth eye movement signal to code depth-sign from motion parallax. *Neuron*, 63(4), 523–532. <https://doi.org/10.1016/j.neuron.2009.07.029>
- Naji, J. J., & Freeman, T. C. (2004). Perceiving depth order during pursuit eye movement. *Vision Research*, 44(26), 3025–3034. <https://doi.org/10.1016/j.visres.2004.07.007>
- Nauta, W. J. H., & Kuypers, H. G. J. M. (1958). Some ascending pathways in the brain stem reticular formation. In *Reticular formation of the brain* (pp. 3–30). Little, Brown.
- Nawrot, M. (2003). Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research*, 43(14), 1553–1562. [https://doi.org/10.1016/S0042-6989\(03\)00144-5](https://doi.org/10.1016/S0042-6989(03)00144-5)

- Nawrot, M., & Joyce, L. (2006). The pursuit theory of motion parallax. *Vision Research*, 46(28), 4709–4725. <https://doi.org/10.1016/j.visres.2006.07.006>
- Niell, C. M. (2015). Cell Types, Circuits, and Receptive Fields in the Mouse Visual Cortex [Publisher: Annual Reviews]. *Annual Review of Neuroscience*, 38(Volume 38, 2015), 413–431. <https://doi.org/10.1146/annurev-neuro-071714-033807>
- Niell, C. M., & Scanziani, M. (2021). How Cortical Circuits Implement Cortical Computations: Mouse Visual Cortex as a Model. *Annual review of neuroscience*, 44, 517. <https://doi.org/10.1146/annurev-neuro-102320-085825>
- Niell, C. M., & Stryker, M. P. (2008). Highly Selective Receptive Fields in Mouse Visual Cortex. *The Journal of Neuroscience*, 28(30), 7520–7536. <https://doi.org/10.1523/JNEUROSCI.0623-08.2008>
- Niell, C. M., & Stryker, M. P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron*, 65(4), 472–479. <https://doi.org/10.1016/j.neuron.2010.01.033>
- Nikara, T., Bishop, P. O., & Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. *Experimental Brain Research*, 6(4), 353–372. <https://doi.org/10.1007/BF00233184>
- Noga, B. R., & Whelan, P. J. (2022). The Mesencephalic Locomotor Region: Beyond Locomotor Control [Publisher: Frontiers]. *Frontiers in Neural Circuits*, 16. <https://doi.org/10.3389/fncir.2022.884785>
- Nurminen, L., Merlin, S., Bijanzadeh, M., Federer, F., & Angelucci, A. (2018). Top-down feedback controls spatial summation and response amplitude in primate visual cortex [Publisher: Nature Publishing Group]. *Nature Communications*, 9(1), 2281. <https://doi.org/10.1038/s41467-018-04500-5>
- Obermayer, K., & Blasdel, G. (1993). Geometry of orientation and ocular dominance columns in monkey striate cortex. *The Journal of Neuroscience*, 13(10), 4114–4129. <https://doi.org/10.1523/JNEUROSCI.13-10-04114.1993>
- O'Connor, D. H., Huber, D., & Svoboda, K. (2009). Reverse engineering the mouse brain [Publisher: Nature Publishing Group]. *Nature*, 461(7266), 923–929. <https://doi.org/10.1038/nature08539>
- Ogle, K. N. (1953). Precision and Validity of Stereoscopic Depth Perception from Double Images* [Publisher: Optica Publishing Group]. *Journal of the Optical Society of America*, 43(10), 906. <https://doi.org/10.1364/josa.43.000906>
- Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P., & Reid, R. C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex [Publisher: Nature Publishing Group UK London]. *Nature*, 433(7026), 597–603.
- Ohki, K., & Reid, R. C. (2007). Specificity and randomness in the visual cortex [Publisher: Elsevier]. *Current opinion in neurobiology*, 17(4), 401–407.
- O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental Neurology*, 51(1), 78–109. [https://doi.org/10.1016/0014-4886\(76\)90055-8](https://doi.org/10.1016/0014-4886(76)90055-8)

- O'Keefe, J., & Burgess, N. (1996). Geometric determinants of the place fields of hippocampal neurons [Publisher: Nature Publishing Group]. *Nature*, *381*(6581), 425–428. <https://doi.org/10.1038/381425a0>
- O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus*, *3*(3), 317–330. <https://doi.org/10.1002/hipo.450030307>
- Olavarria, J., & Montero, V. M. (1989). Organization of visual cortex in the mouse revealed by correlating callosal and striate-extrastriate connections. *Visual Neuroscience*, *3*(1), 59–69. <https://doi.org/10.1017/s0952523800012517>
- Olson, C. R. (1980). Spatial localization in cats reared with strabismus. *Journal of Neurophysiology*, *43*(3), 792–806. <https://doi.org/10.1152/jn.1980.43.3.792>
- Ono, H., Rogers, B. J., Ohmi, M., & Ono, M. E. (1988). Dynamic Occlusion and Motion Parallax in Depth Perception. *Perception*, *17*(2), 255–266. <https://doi.org/10.1068/p170255>
- Orsolio, I., Rio, M., Mšic-Flogel, T. D., & Znamenskiy, P. (2021). Mesoscale cortical dynamics reflect the interaction of sensory evidence and temporal expectation during perceptual decision-making [Publisher: Elsevier]. *Neuron*, *109*(11), 1861–1875.e10. <https://doi.org/10.1016/j.neuron.2021.03.031>
- Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L. F., Dalgleish, H., Carandini, M., & Harris, K. D. (2017, July). Suite2p: Beyond 10,000 neurons with standard two-photon microscopy [Pages: 061507 Section: New Results]. <https://doi.org/10.1101/061507>
- Packer, A. M., & Yuste, R. (2011). Dense, Unspecific Connectivity of Neocortical Parvalbumin-Positive Interneurons: A Canonical Microcircuit for Inhibition? [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, *31*(37), 13260–13271. <https://doi.org/10.1523/JNEUROSCI.3131-11.2011>
- Packwood, J., & Gordon, B. (1975). Stereopsis in normal domestic cat, Siamese cat, and cat raised with alternating monocular occlusion [Publisher: American Physiological Society]. *Journal of Neurophysiology*, *38*(6), 1485–1499. <https://doi.org/10.1152/jn.1975.38.6.1485>
- Pakan, J. M., Lowe, S. C., Dylida, E., Keemink, S. W., Currie, S. P., Coutts, C. A., & Rochefort, N. L. (2016). Behavioral-state modulation of inhibition is context-dependent and cell type specific in mouse visual cortex (T. D. Mšic-Flogel, Ed.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, *5*, e14985. <https://doi.org/10.7554/eLife.14985>
- Panda, S., Sato, T. K., Castrucci, A. M., Rollag, M. D., DeGrip, W. J., Hogenesch, J. B., Provencio, I., & Kay, S. A. (2002). Melanopsin (Opn4) Requirement for Normal Light-Induced Circadian Phase Shifting [Publisher: American Association for the Advancement of Science]. *Science*, *298*(5601), 2213–2216. <https://doi.org/10.1126/science.1076848>

- Paquette, D., & Bigras, M. (2018). Duration of Childhood. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1–9). Springer International Publishing. https://doi.org/10.1007/978-3-319-16999-6_2384-1
- Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, 8(5), 379–391.
- Parker, P. R. L., Abe, E. T. T., Leonard, E. S. P., Martins, D. M., & Niell, C. M. (2022). Joint coding of visual input and eye/head position in V1 of freely moving mice. *Neuron*, 110(23), 3897–3906.e5. <https://doi.org/10.1016/j.neuron.2022.08.029>
- Parker, P. R., Abe, E. T., Beatie, N. T., Leonard, E. S., Martins, D. M., Sharp, S. L., Wyrick, D. G., Mazzucato, L., & Niell, C. M. (2022). Distance estimation from monocular cues in an ethological visuomotor task (N. J. Cowan, T. Moore, & N. J. Cowan, Eds.) [Publisher: eLife Sciences Publications, Ltd]. *eLife*, 11, e74708. <https://doi.org/10.7554/eLife.74708>
- Pasternak, T., Tompkins, J., & Olson, C. (1995). The role of striate cortex in visual function of the cat. *The Journal of Neuroscience*, 15(3), 1940–1950. <https://doi.org/10.1523/JNEUROSCI.15-03-01940.1995>
- Patterson, G. H., & Piston, D. W. (2000). Photobleaching in Two-Photon Excitation Microscopy. *Biophysical Journal*, 78(4), 2159–2162. [https://doi.org/10.1016/S0006-3495\(00\)76762-2](https://doi.org/10.1016/S0006-3495(00)76762-2)
- Peterhans, E., & von der Heydt, R. (1993). Functional Organization of Area V2 in the Alert Macaque [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-9568.1993.tb00517.x>]. *European Journal of Neuroscience*, 5(5), 509–524. <https://doi.org/10.1111/j.1460-9568.1993.tb00517.x>
- Petersen, C. C. H., & Crochet, S. (2013). Synaptic Computation and Sensory Processing in Neocortical Layer 2/3 [Publisher: Elsevier]. *Neuron*, 78(1), 28–48. <https://doi.org/10.1016/j.neuron.2013.03.020>
- Petros, T. J., Rebsam, A., & Mason, C. A. (2008). Retinal Axon Growth at the Optic Chiasm: To Cross or Not to Cross. *Annual Review of Neuroscience*, 31(1), 295–315. <https://doi.org/10.1146/annurev.neuro.31.060407.125609>
- Pettigrew, J., Nikara, T., & Bishop, P. (1968). Binocular interaction on single units in cat striate cortex: Simultaneous stimulation by single moving slit with receptive fields in correspondence [Publisher: Springer]. *Experimental brain research*, 6, 391–410.
- Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., & Scanziani, M. (2013). Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons [Publisher: Nature Publishing Group]. *Nature Neuroscience*, 16(8), 1068–1076. <https://doi.org/10.1038/nn.3446>
- Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey [Publisher: American Physiolog-

- ical Society]. *Journal of Neurophysiology*, 40(6), 1392–1405. <https://doi.org/10.1152/jn.1977.40.6.1392>
- Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 8(12), 4531–4550. <https://doi.org/10.1523/JNEUROSCI.08-12-04531.1988>
- Poggio, G. F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, 25(3), 397–406. [https://doi.org/10.1016/0042-6989\(85\)90065-3](https://doi.org/10.1016/0042-6989(85)90065-3)
- Poggio, G. F., & Poggio, T. (1984). The Analysis of Stereopsis [Publisher: Annual Reviews]. *Annual Review of Neuroscience*, 7(Volume 7, 1984), 379–412. <https://doi.org/10.1146/annurev.ne.07.030184.002115>
- Poggio, G. F., & Talbot, W. H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1981.sp013759>]. *The Journal of Physiology*, 315(1), 469–492. <https://doi.org/10.1113/jphysiol.1981.sp013759>
- Polack, P.-O., Friedman, J., & Golshani, P. (2013). Cellular mechanisms of brain state-dependent gain modulation in visual cortex [Publisher: Nature Publishing Group]. *Nature Neuroscience*, 16(9), 1331–1339. <https://doi.org/10.1038/nn.3464>
- Pouget, A. (1993). Egocentric Spaw Representation in Early Viion. *Journal of Cognitive Neuroscience*, 5(2).
- Pouget, A., & Sejnowski, T. J. (1997). Spatial Transformations in the Parietal Cortex Using Basis Functions. *Journal of Cognitive Neuroscience*, 9(2), 222–237. <https://doi.org/10.1162/jocn.1997.9.2.222>
- Priebe, N. J., & Ferster, D. (2008). Inhibition, Spike Threshold, and Stimulus Selectivity in Primary Visual Cortex [Publisher: Elsevier]. *Neuron*, 57(4), 482–497. <https://doi.org/10.1016/j.neuron.2008.02.005>
- Priebe, N. J., Lisberger, S. G., & Movshon, J. A. (2006). Tuning for Spatiotemporal Frequency and Speed in Directionally Selective Neurons of Macaque Striate Cortex [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 26(11), 2941–2950. <https://doi.org/10.1523/JNEUROSCI.3936-05.2006>
- Ptito, M., Lepore, F., & Guillemot, J.-P. (1992). Loss of stereopsis following lesions of cortical areas 17/18 in the cat. *Experimental Brain Research*, 89(3), 521–530. <https://doi.org/10.1007/BF00229877>
- Qian, N. (1997). Binocular Disparity and the Perception of Depth [Publisher: Elsevier]. *Neuron*, 18(3), 359–368. [https://doi.org/10.1016/S0896-6273\(00\)81238-6](https://doi.org/10.1016/S0896-6273(00)81238-6)
- Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79–87. <https://doi.org/10.1038/4580>

- Read, J. C. (2021). Binocular vision and stereopsis across the animal kingdom [Publisher: Annual Reviews]. *Annual Review of Vision Science*, 7(1), 389–415.
- Reese, B. E. (1988). 'Hidden lamination' in the dorsal lateral geniculate nucleus: The functional organization of this thalamic region in the rat. *Brain Research*, 472(2), 119–137. [https://doi.org/10.1016/0165-0173\(88\)90017-3](https://doi.org/10.1016/0165-0173(88)90017-3)
- Reese, B. E., & Jeffery, G. (1983). Crossed and uncrossed visual topography in dorsal lateral geniculate nucleus of the pigmented rat [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 49(4), 877–885. <https://doi.org/10.1152/jn.1983.49.4.877>
- Reichelt, S., Häussler, R., Fütterer, G., & Leister, N. (2010). Depth cues in human visual perception and their realization in 3D displays. *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV*, 7690, 92–103.
- Reid, R. C., & Alonso, J.-M. (1995). Specificity of monosynaptic connections from thalamus to visual cortex. *Nature*, 378(6554), 281–284. <https://doi.org/10.1038/378281a0>
- Reimer, J., Froudarakis, E., Cadwell, C. R., Yatsenko, D., Denfield, G. H., & Tolias, A. S. (2014). Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness [Publisher: Elsevier]. *Neuron*, 84(2), 355–362. <https://doi.org/10.1016/j.neuron.2014.09.033>
- Ringach, D. L., Mineault, P. J., Tring, E., Olivas, N. D., Garcia-Junco-Clemente, P., & Trachtenberg, J. T. (2016). Spatial clustering of tuning in mouse primary visual cortex [Publisher: Nature Publishing Group UK London]. *Nature communications*, 7(1), 12270.
- Roe, A., & Ts'o, D. (1995). Visual topography in primate V2: Multiple representation across functional stripes. *The Journal of Neuroscience*, 15(5), 3689–3715. <https://doi.org/10.1523/JNEUROSCI.15-05-03689.1995>
- Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, 8(2), 125–134.
- Rogers, S., & Rogers, B. J. (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception & Psychophysics*, 52(4), 446–452. <https://doi.org/10.3758/BF03206704>
- Román Rosón, M., Bauer, Y., Kotkat, A. H., Berens, P., Euler, T., & Busse, L. (2019). Mouse dLGN Receives Functional Input from a Diverse Population of Retinal Ganglion Cells with Limited Convergence. *Neuron*, 102(2), 462–476.e8. <https://doi.org/10.1016/j.neuron.2019.01.040>
- Rosenbluth, D., & Allman, J. M. (2002). The Effect of Gaze Angle and Fixation Distance on the Responses of Neurons in V1, V2, and V4 [Publisher: Elsevier]. *Neuron*, 33(1), 143–149. [https://doi.org/10.1016/S0896-6273\(01\)00559-1](https://doi.org/10.1016/S0896-6273(01)00559-1)

- Roth, M. M., Helmchen, F., & Kampa, B. M. (2012). Distinct Functional Properties of Primary and Posteromedial Visual Area of Mouse Neocortex. *Journal of Neuroscience*, 32(28), 9716–9726. <https://doi.org/10.1523/JNEUROSCI.0110-12.2012>
- Roth, M. M., Dahmen, J. C., Muir, D. R., Imhof, F., Martini, F. J., & Hofer, S. B. (2016). Thalamic nuclei convey diverse contextual information to layer 1 of visual cortex [Publisher: Nature Publishing Group US New York]. *Nature neuroscience*, 19(2), 299–307.
- Royden, C. S., Baker, J. F., & Allman, J. (1988). Perceptions of Depth Elicited by Occluded and Shearing Motions of Random Dots [Publisher: SAGE Publications Ltd STM]. *Perception*, 17(3), 289–296. <https://doi.org/10.1068/p170289>
- Ruff, D. A., & Cohen, M. R. (2014). Attention can either increase or decrease spike count correlations in visual cortex [Publisher: Nature Publishing Group]. *Nature Neuroscience*, 17(11), 1591–1597. <https://doi.org/10.1038/nn.3835>
- Russell, J. T. (1932). Depth Discrimination in the Rat. *The Pedagogical Seminary and Journal of Genetic Psychology*, 40(1), 136–161. <https://doi.org/10.1080/08856559.1932.10534211>
- Saito, H., 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. <https://doi.org/10.1523/JNEUROSCI.06-01-00145.1986>
- Sakata, H., Shibutani, H., & Kawano, K. (1980). Spatial properties of visual fixation neurons in posterior parietal association cortex of the monkey. *Journal of Neurophysiology*, 43(6), 1654–1672. <https://doi.org/10.1152/jn.1980.43.6.1654>
- Saleem, A. B., Ayaz, A., Jeffery, K. J., Harris, K. D., & Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. *Nature Neuroscience*, 16(12), 1864–1869. <https://doi.org/10.1038/nn.3567>
- Saleem, A. B., Diamanti, E. M., Fournier, J., Harris, K. D., & Carandini, M. (2018). Coherent encoding of subjective spatial position in visual cortex and hippocampus. *Nature*, 562(7725), 124–127.
- Salinas, E., & Abbott, L. F. (1996). A model of multiplicative neural responses in parietal cortex. *Proceedings of the National Academy of Sciences*, 93(21), 11956–11961. <https://doi.org/10.1073/pnas.93.21.11956>
- Samonds, J. M., Choi, V., & Priebe, N. J. (2019). Mice discriminate stereoscopic surfaces without fixating in depth. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 39(41), 8024–8037. <https://doi.org/10.1523/jneurosci.0895-19.2019>
- Saravanan, V., Berman, G. J., & Sober, S. J. (2020). Application of the hierarchical bootstrap to multi-level data in neuroscience. *Neurons, behavior, data analysis and theory*, 3(5). Retrieved January 25, 2025, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7906290/>

- Sattler, N. J., & Wehr, M. (2021). A Head-Mounted Multi-Camera System for Electrophysiology and Behavior in Freely-Moving Mice [Publisher: Frontiers]. *Frontiers in Neuroscience*, 14. <https://doi.org/10.3389/fnins.2020.592417>
- Savelli, F., Yoganarasimha, D., & Knierim, J. J. (2008). Influence of boundary removal on the spatial representations of the medial entorhinal cortex. *Hippocampus*, 18(12), 1270–1282. <https://doi.org/10.1002/hipo.20511>
- Savier, E. L., Chen, H., & Cang, J. (2019). Effects of locomotion on visual responses in the mouse superior colliculus [Publisher: Soc Neuroscience]. *Journal of Neuroscience*, 39(47), 9360–9368.
- Sayed, M., Gibson, J., Watson, J., Prisacariu, V., Firman, M., & Godard, C. (2022). SimpleRecon: 3D Reconstruction Without 3D Convolutions. In S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, & T. Hassner (Eds.), *Computer Vision ECCV 2022* (pp. 1–19). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-19827-4_1
- Schiller, P. H., Slocum, W. M., Jao, B., & Weiner, V. S. (2011). The integration of disparity, shading and motion parallax cues for depth perception in humans and monkeys. *Brain research*, 1377, 67–77. <https://doi.org/10.1016/j.brainres.2011.01.003>
- Schmidt, T. M., & Kofuji, P. (2008). Novel insights into non-image forming visual processing in the retina. *Cellscience*, 5(1), 77–83. Retrieved November 26, 2024, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2890289/>
- Scholl, B., Burge, J., & Priebe, N. J. (2013). Binocular integration and disparity selectivity in mouse primary visual cortex. *Journal of Neurophysiology*, 109(12), 3013–3024. <https://doi.org/10.1152/jn.01021.2012>
- Scholtyssek, C., Dacke, M., Kröger, R., & Baird, E. (2014). Control of self-motion in dynamic fluids: Fish do it differently from bees. *Biology Letters*, 10(5), 20140279. <https://doi.org/10.1098/rsbl.2014.0279>
- Schor, C. M., Maxwell, J. S., McCANDLESS, J., & Graf, E. (2002). Adaptive Control of Vergence in Humans [eprint: <https://nyaspubs.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1749-6632.2002.tb02828.x>]. *Annals of the New York Academy of Sciences*, 956(1), 297–305. <https://doi.org/10.1111/j.1749-6632.2002.tb02828.x>
- Schuett, S., Bonhoeffer, T., & Hübener, M. (2002). Mapping Retinotopic Structure in Mouse Visual Cortex with Optical Imaging [Publisher: Society for Neuroscience Section: ARTICLE]. *Journal of Neuroscience*, 22(15), 6549–6559. <https://doi.org/10.1523/JNEUROSCI.22-15-06549.2002>
- Seabrook, T. A., Burbridge, T. J., Crair, M. C., & Huberman, A. D. (2017). Architecture, Function, and Assembly of the Mouse Visual System. *Annual Review of Neuroscience*, 40(1), 499–538. <https://doi.org/10.1146/annurev-neuro-071714-033842>
- Shang, C., Liu, A., Li, D., Xie, Z., Chen, Z., Huang, M., Li, Y., Wang, Y., Shen, W. L., & Cao, P. (2019). A subcortical excitatory circuit for sensory-triggered predatory

- hunting in mice [Publisher: Nature Publishing Group]. *Nature Neuroscience*, 22(6), 909–920. <https://doi.org/10.1038/s41593-019-0405-4>
- Shang, C., Liu, Z., Chen, Z., Shi, Y., Wang, Q., Liu, S., Li, D., & Cao, P. (2015). A parvalbumin-positive excitatory visual pathway to trigger fear responses in mice [Publisher: American Association for the Advancement of Science]. *Science*, 348(6242), 1472–1477. <https://doi.org/10.1126/science.aaa8694>
- Shang, W., Xie, S., Feng, W., Li, Z., Jia, J., Cao, X., Shen, Y., Li, J., Shi, H., Gu, Y., Weng, S.-J., Lin, L., Pan, Y.-H., & Yuan, X.-B. (2024). A non-image-forming visual circuit mediates the innate fear of heights in male mice [Publisher: Nature Publishing Group]. *Nature Communications*, 15(1), 3746. <https://doi.org/10.1038/s41467-024-48147-x>
- Shen, J., & Colonnese, M. T. (2016). Development of Activity in the Mouse Visual Cortex. *The Journal of Neuroscience*, 36(48), 12259–12275. <https://doi.org/10.1523/jneurosci.1903-16.2016>
- Sherrington, C. S. (1906). Observations on the scratch-reflex in the spinal dog. *The Journal of Physiology*, 34(1-2), 1–50. Retrieved December 24, 2024, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1465804/>
- Shik, M. L., & Orlovsky, G. N. (1976). Neurophysiology of locomotor automatism [Publisher: American Physiological Society]. *Physiological Reviews*, 56(3), 465–501. <https://doi.org/10.1152/physrev.1976.56.3.465>
- Shimono, K., Tam, W. J., Stelmach, L., & Hildreth, E. (2002). Stereoillusive motion concomitant with lateral head movements. *Perception & Psychophysics*, 64(8), 1218–1226. <https://doi.org/10.3758/BF03194767>
- Sitko, A. A., Kuwajima, T., & Mason, C. (2018). Eye-specific segregation and differential fasciculation of developing retinal ganglion cell axons in the mouse visual pathway. *The Journal of comparative neurology*, 526(7), 1077–1096. <https://doi.org/10.1002/cne.24392>
- Skottun, B. C., De Valois, R. L., Grosof, D. H., Movshon, J. A., Albrecht, D. G., & Bonds, A. B. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision Research*, 31(7), 1078–1086. [https://doi.org/10.1016/0042-6989\(91\)90033-2](https://doi.org/10.1016/0042-6989(91)90033-2)
- Skyberg, R. J., & Niell, C. M. (2024). Natural visual behavior and active sensing in the mouse. *Current Opinion in Neurobiology*, 86, 102882. <https://doi.org/10.1016/j.conb.2024.102882>
- Smyth, D., Willmore, B., Baker, G. E., Thompson, I. D., & Tolhurst, D. J. (2003). The Receptive-Field Organization of Simple Cells in Primary Visual Cortex of Ferrets under Natural Scene Stimulation. *The Journal of Neuroscience*, 23(11), 4746–4759. <https://doi.org/10.1523/JNEUROSCI.23-11-04746.2003>
- Sobel, E. C. (1990a). Depth perception by motion parallax and paradoxical parallax in the locust. *Naturwissenschaften*, 77(5), 241–243. <https://doi.org/10.1007/BF01138494>

- Sobel, E. (1990b). The locust's use of motion parallax to measure distance. *Journal of Comparative Physiology A*, 167(5). <https://doi.org/10.1007/BF00192653>
- Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B., & Moser, E. I. (2008). Representation of Geometric Borders in the Entorhinal Cortex [Publisher: American Association for the Advancement of Science]. *Science*, 322(5909), 1865–1868. <https://doi.org/10.1126/science.1166466>
- Spalding, D. A. (1875). Instinct and Acquisition*. *Nature*, 12(310), 507–508. <https://doi.org/10.1038/012507a0>
- Sparks, D. L., Lee, C., & Rohrer, W. H. (1990). Population Coding of the Direction, Amplitude, and Velocity of Saccadic Eye Movements by Neurons in the Superior Colliculus [Publisher: Cold Spring Harbor Laboratory Press]. *Cold Spring Harbor Symposia on Quantitative Biology*, 55, 805–811. <https://doi.org/10.1101/SQB.1990.055.01.075>
- Sparks, D. L. (2002). The brainstem control of saccadic eye movements [Publisher: Nature Publishing Group]. *Nature Reviews Neuroscience*, 3(12), 952–964. <https://doi.org/10.1038/nrn986>
- Spector, R. H. (1990). Visual Fields. In H. K. Walker, W. D. Hall, & J. W. Hurst (Eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations* (3rd). Butterworths. Retrieved November 19, 2024, from <http://www.ncbi.nlm.nih.gov/books/NBK220/>
- Spillmann, L. (2014). Receptive Fields of Visual Neurons: The Early Years [Publisher: SAGE Publications Ltd STM]. *Perception*, 43(11), 1145–1176. <https://doi.org/10.1068/p7721>
- Spratling, M. W. (2010). Predictive Coding as a Model of Response Properties in Cortical Area V1 [Publisher: Society for Neuroscience Section: Articles]. *Journal of Neuroscience*, 30(9), 3531–3543. <https://doi.org/10.1523/JNEUROSCI.4911-09.2010>
- Srinivasan, M. V., Zhang, S. W., Chahl, J. S., Barth, E., & Venkatesh, S. (2000). How honeybees make grazing landings on flat surfaces. *Biological Cybernetics*, 83, 171–183.
- Srinivasan, M. V. (1992). Distance Perception in Insects [Publisher: [Association for Psychological Science, Sage Publications, Inc.]]. *Current Directions in Psychological Science*, 1(1), 22–26. Retrieved November 21, 2024, from <https://www.jstor.org/stable/20182119>
- Stahl, J. S. (2004). Using eye movements to assess brain function in mice. *Vision Research*, 44(28), 3401–3410. <https://doi.org/10.1016/j.visres.2004.09.011>
- Steinmetz, N. A., Aydin, C., Lebedeva, A., Okun, M., Pachitariu, M., Bauza, M., Beau, M., Bhagat, J., Böhm, C., Broux, M., Chen, S., Colonell, J., Gardner, R. J., Karsh, B., Kloosterman, F., Kostadinov, D., Mora-Lopez, C., OCallaghan, J., Park, J., ... Harris, T. D. (2021). Neuropixels 2.0: A miniaturized high-density probe for

- stable, long-term brain recordings. *Science (New York, N.Y.)*, 372(6539), eabf4588. <https://doi.org/10.1126/science.abf4588>
- Straschill, M., & Hoffman, K. P. (1969). Response characteristics of movement-detecting neurons in pretectal region of the cat [Place: Netherlands Publisher: Elsevier Science]. *Experimental Neurology*, 25(2), 165–176. [https://doi.org/10.1016/0014-4886\(69\)90041-7](https://doi.org/10.1016/0014-4886(69)90041-7)
- Stringer, C., Pachitariu, M., Steinmetz, N., Reddy, C. B., Carandini, M., & Harris, K. D. (2019). Spontaneous behaviors drive multidimensional, brainwide activity. *Science*, 364(6437), eaav7893. <https://doi.org/10.1126/science.aav7893>
- Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: A generalist algorithm for cellular segmentation [Publisher: Nature Publishing Group]. *Nature Methods*, 18(1), 100–106. <https://doi.org/10.1038/s41592-020-01018-x>
- Strother, J. A., Wu, S.-T., Wong, A. M., Nern, A., Rogers, E. M., Le, J. Q., Rubin, G. M., & Reiser, M. B. (2017). The Emergence of Directional Selectivity in the Visual Motion Pathway of Drosophila. *Neuron*, 94(1), 168–182.e10. <https://doi.org/10.1016/j.neuron.2017.03.010>
- Sweet, B., & Kaiser, M. (2011). Depth Perception, Cueing, and Control. *AIAA Modeling and Simulation Technologies Conference*. <https://doi.org/10.2514/6.2011-6424>
- Syeda, A., Zhong, L., Tung, R., Long, W., Pachitariu, M., & Stringer, C. (2024). Facemap: A framework for modeling neural activity based on orofacial tracking [Publisher: Nature Publishing Group]. *Nature Neuroscience*, 27(1), 187–195. <https://doi.org/10.1038/s41593-023-01490-6>
- Takemura, A., Inoue, Y., Kawano, K., Quaia, C., & Miles, F. A. (2001). Single-Unit Activity in Cortical Area MST Associated With Disparity-Vergence Eye Movements: Evidence for Population Coding [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 85(5), 2245–2266. <https://doi.org/10.1152/jn.2001.85.5.2245>
- Talluri, B. C., Kang, I., Lazere, A., Quinn, K. R., Kaliss, N., Yates, J. L., Butts, D. A., & Nienborg, H. (2023). Activity in primate visual cortex is minimally driven by spontaneous movements [Publisher: Nature Publishing Group US New York]. *Nature Neuroscience*, 26(11), 1953–1959.
- Tan, L., Ringach, D. L., & Trachtenberg, J. T. (2022). The Development of Receptive Field Tuning Properties in Mouse Binocular Primary Visual Cortex. *The Journal of Neuroscience*, 42(17), 3546–3556. <https://doi.org/10.1523/jneurosci.1702-21.2022>
- Tan, L., Ringach, D. L., Zipursky, S. L., & Trachtenberg, J. T. (2021). Vision is required for the formation of binocular neurons prior to the classical critical period [Publisher: Elsevier]. *Current Biology*, 31(19), 4305–4313.e5. <https://doi.org/10.1016/j.cub.2021.07.053>

- Tanabe, S., Umeda, K., & Fujita, I. (2004). Rejection of False Matches for Binocular Correspondence in Macaque Visual Cortical Area V4. *The Journal of Neuroscience*, 24(37), 8170–8180. <https://doi.org/10.1523/JNEUROSCI.5292-03.2004>
- Thompson, A., Gribizis, A., Chen, C., & Crair, M. C. (2017). Activity-dependent development of visual receptive fields. *Current Opinion in Neurobiology*, 42, 136–143. <https://doi.org/10.1016/j.conb.2016.12.007>
- Thorndike, E. (1899). The instinctive reaction of young chicks. *Psychological Review*, 6(3), 282–291. <https://doi.org/10.1037/h0073000>
- Tian, N., & Copenhagen, D. R. (2001). Visual Deprivation Alters Development of Synaptic Function in Inner Retina after Eye Opening [Publisher: Elsevier]. *Neuron*, 32(3), 439–449. [https://doi.org/10.1016/S0896-6273\(01\)00470-6](https://doi.org/10.1016/S0896-6273(01)00470-6)
- Tohmi, M., Meguro, R., Tsukano, H., Hishida, R., & Shibuki, K. (2014). The extrageniculate visual pathway generates distinct response properties in the higher visual areas of mice. *Current biology: CB*, 24(6), 587–597. <https://doi.org/10.1016/j.cub.2014.01.061>
- Tootell, R., Switkes, E., Silverman, M., & Hamilton, S. (1988). Functional anatomy of macaque striate cortex. II. Retinotopic organization. *The Journal of Neuroscience*, 8(5), 1531–1568. <https://doi.org/10.1523/JNEUROSCI.08-05-01531.1988>
- Tovée, M. J. (2008). *An introduction to the visual system*. Cambridge University Press.
- Triplett, J. W. (2014). Molecular guidance of retinotopic map development in the mid-brain. *Current Opinion in Neurobiology*, 24, 7–12. <https://doi.org/10.1016/j.conb.2013.07.006>
- Triplett, J. W., Phan, A., Yamada, J., & Feldheim, D. A. (2012). Alignment of multimodal sensory input in the superior colliculus through a gradient-matching mechanism [Publisher: Soc Neuroscience]. *Journal of Neuroscience*, 32(15), 5264–5271.
- Trotter, Y., & Celebrini, S. (1999). Gaze direction controls response gain in primary visual-cortex neurons [Publisher: Nature Publishing Group]. *Nature*, 398(6724), 239–242. <https://doi.org/10.1038/18444>
- Trychin, S., & Walk, R. D. (1964). A study of the depth perception of monocular hooded rats on the visual cliff. *Psychonomic Science*, 1(1-12), 53–54. <https://doi.org/10.3758/bf03342786>
- Tsuruhara, A., Corrow, S., Kanazawa, S., Yamaguchi, M. K., & Yonas, A. (2014). Infants ability to respond to depth from the retinal size of human faces: Comparing monocular and binocular preferential-looking [Publisher: Elsevier]. *Infant Behavior and Development*, 37(4), 562–570.
- Turnbull, O. H., Driver, J., & McCarthy, R. A. (2004). 2D But Not 3D: Pictorial-Depth Deficits in a Case of Visual Agnosia. *Cortex*, 40(4), 723–738. [https://doi.org/10.1016/S0010-9452\(08\)70167-9](https://doi.org/10.1016/S0010-9452(08)70167-9)
- Turski, J. (2023). The horopter: Old and new [Publisher: SAGE Publications Ltd STM]. *Perception*, 52(6), 412–422. <https://doi.org/10.1177/03010066231170380>

- van Beest, E. H., Mukherjee, S., Kirchberger, L., Schnabel, U. H., van der Togt, C., Teeuwen, R. R. M., Barsegyan, A., Meyer, A. F., Poort, J., Roelfsema, P. R., & Self, M. W. (2021). Mouse visual cortex contains a region of enhanced spatial resolution [Publisher: Nature Publishing Group]. *Nature Communications*, 12(1), 4029. <https://doi.org/10.1038/s41467-021-24311-5>
- Van den Bergh, G., Zhang, B., Arckens, L., & Chino, Y. M. (2010). Receptive-field properties of V1 and V2 neurons in mice and macaque monkeys. *The Journal of Comparative Neurology*, 518(11), 2051–2070. <https://doi.org/10.1002/cne.22321>
- Van Hooser, S. D., Heimel, J. A., & Nelson, S. B. (2005). Functional cell classes and functional architecture in the early visual system of a highly visual rodent. *Progress in Brain Research*, 149, 127–145. [https://doi.org/10.1016/S0079-6123\(05\)49010-X](https://doi.org/10.1016/S0079-6123(05)49010-X)
- Vangeneugden, J., Beest, E. H. v., Cohen, M. X., Lorteije, J. A. M., Mukherjee, S., Kirchberger, L., Montijn, J. S., Thamizharasu, P., Camillo, D., Levelt, C. N., Roelfsema, P. R., Self, M. W., & Heimel, J. A. (2019). Activity in Lateral Visual Areas Contributes to Surround Suppression in Awake Mouse V1 [Publisher: Elsevier]. *Current Biology*, 29(24), 4268–4275.e7. <https://doi.org/10.1016/j.cub.2019.10.037>
- Vélez-Fort, M., Bracey, E. F., Keshavarzi, S., Rousseau, C. V., Cossell, L., Lenzi, S. C., Strom, M., & Margrie, T. W. (2018). A circuit for integration of head-and visual-motion signals in layer 6 of mouse primary visual cortex. *Neuron*, 98(1), 179–191.
- Verhoef, B.-E., Vogels, R., & Janssen, P. (2010). Contribution of Inferior Temporal and Posterior Parietal Activity to Three-Dimensional Shape Perception [Publisher: Elsevier]. *Current Biology*, 20(10), 909–913. <https://doi.org/10.1016/j.cub.2010.03.058>
- Vinck, M., Batista-Brito, R., Knoblich, U., & Cardin, J. A. (2015). Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. *Neuron*, 86(3), 740–754. <https://doi.org/10.1016/j.neuron.2015.03.028>
- von der Heydt, R., Adorjani, C., Hännny, P., & Baumgartner, G. (1978). Disparity sensitivity and receptive field incongruity of units in the cat striate cortex. *Experimental Brain Research*, 31(4), 523–545. <https://doi.org/10.1007/BF00239810>
- Vong, L., Ye, C., Yang, Z., Choi, B., Chua, S., & Lowell, B. B. (2011). Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons [Publisher: Elsevier]. *Neuron*, 71(1), 142–154.
- Wade, N. J. (2002). Charles Wheatstone (1802–1875) [Publisher: SAGE Publications Ltd STM]. *Perception*, 31(3), 265–272. <https://doi.org/10.1068/p3103ed>
- Wagor, E., Mangini, N. J., & Pearlman, A. L. (1980). Retinotopic organization of striate and extrastriate visual cortex in the mouse. *Journal of Comparative Neurology*, 193(1), 187–202. <https://doi.org/10.1002/cne.901930113>
- Walk, R. D., Gibson, E. J., & Tighe, T. J. (1957). Behavior of light- and dark-reared rats on a visual cliff. *Science (New York, N. Y.)*, 126(3263), 80–81. <https://doi.org/10.1126/science.126.3263.80-a>

- Walk, R. D., & Dodge, S. H. (1962). Visual depth perception of a 10-month-old monocular human infant. *Science*, *137*(3529), 529–530. <https://doi.org/10.1126/science.137.3529.529>
- Walk, R. D., & Gibson, E. J. (1961). A comparative and analytical study of visual depth perception. *Psychological Monographs: General and Applied*, *75*(15), 1.
- Walk, R. D., Trychin, S., & Karmel, B. Z. (1965). Depth perception in the dark-reared rat as a function of time in the dark. *Psychonomic Science*, *3*(1-12), 9–10. <https://doi.org/10.3758/bf03342991>
- Walker, E. Y., Sinz, F. H., Cobos, E., Muhammad, T., Froudarakis, E., Fahey, P. G., Ecker, A. S., Reimer, J., Pitkow, X., & Tolia, A. S. (2019). Inception loops discover what excites neurons most using deep predictive models [Publisher: Nature Publishing Group]. *Nature Neuroscience*, *22*(12), 2060–2065. <https://doi.org/10.1038/s41593-019-0517-x>
- Wallace, D. J., Greenberg, D. S., Sawinski, J., Rulla, S., Notaro, G., & Kerr, J. N. D. (2013). Rats maintain an overhead binocular field at the expense of constant fusion [Publisher: Nature Publishing Group]. *Nature*, *498*(7452), 65–69. <https://doi.org/10.1038/nature12153>
- Wallace, G. K. (1959). Visual Scanning in the Desert Locust *Schistocerca Gregaria* Forskål. *Journal of Experimental Biology*, *36*(3), 512–525. <https://doi.org/10.1242/jeb.36.3.512>
- Wandell, B. A. (1995). *Foundations of vision* [Pages: xvi, 476]. Sinauer Associates.
- Wandell, B. A., & Winawer, J. (2015). Computational neuroimaging and population receptive fields. *Trends in Cognitive Sciences*, *19*(6), 349–357. <https://doi.org/10.1016/j.tics.2015.03.009>
- Wang, C., Chen, X., & Knierim, J. J. (2020). Egocentric and allocentric representations of space in the rodent brain. *Current Opinion in Neurobiology*, *60*, 12–20. <https://doi.org/10.1016/j.conb.2019.11.005>
- Wang, C., Chen, X., Lee, H., Deshmukh, S. S., Yoganarasimha, D., Savelli, F., & Knierim, J. J. (2018). Egocentric coding of external items in the lateral entorhinal cortex [Publisher: American Association for the Advancement of Science]. *Science*, *362*(6417), 945–949. <https://doi.org/10.1126/science.aau4940>
- Wang, C., & Dreher, B. (1996). Binocular interactions and disparity coding in area 21a of cat extrastriate visual cortex. *Experimental Brain Research*, *108*(2), 257–272. <https://doi.org/10.1007/BF00228099>
- Wang, L., Liu, M., Segreaves, M. A., & Cang, J. (2015). Visual Experience Is Required for the Development of Eye Movement Maps in the Mouse Superior Colliculus [Publisher: Society for Neuroscience Section: Brief Communications]. *Journal of Neuroscience*, *35*(35), 12281–12286. <https://doi.org/10.1523/JNEUROSCI.0117-15.2015>

- Wang, L., Sarnaik, R., Rangarajan, K., Liu, X., & Cang, J. (2010). Visual Receptive Field Properties of Neurons in the Superficial Superior Colliculus of the Mouse. *The Journal of Neuroscience*, *30*(49), 16573–16584. <https://doi.org/10.1523/jneurosci.3305-10.2010>
- Wang, Q., & Burkhalter, A. (2007). Area map of mouse visual cortex [eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/cne.21286>]. *Journal of Comparative Neurology*, *502*(3), 339–357. <https://doi.org/10.1002/cne.21286>
- Wang, Q., Ding, S.-L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., Dolbeare, T., Blanchard, B., Dee, N., Wakeman, W., Hirokawa, K. E., Szafer, A., Sunkin, S. M., Oh, S. W., Bernard, A., ... Ng, L. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. *Cell*, *181*(4), 936–953.e20. <https://doi.org/10.1016/j.cell.2020.04.007>
- Wang, Q., Gao, E., & Burkhalter, A. (2011). Gateways of Ventral and Dorsal Streams in Mouse Visual Cortex. *The Journal of Neuroscience*, *31*(5), 1905. <https://doi.org/10.1523/JNEUROSCI.3488-10.2011>
- Wang, Q., Sporns, O., & Burkhalter, A. (2012). Network analysis of corticocortical connections reveals ventral and dorsal processing streams in mouse visual cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *32*(13), 4386–4399. <https://doi.org/10.1523/JNEUROSCI.6063-11.2012>
- Waters, D. A., & Abulula, H. H. (2007). Using bat-modelled sonar as a navigational tool in virtual environments. *International Journal of Human-Computer Studies*, *65*(10), 873–886. <https://doi.org/10.1016/j.ijhcs.2007.06.001>
- Wei, P., Liu, N., Zhang, Z., Liu, X., Tang, Y., He, X., Wu, B., Zhou, Z., Liu, Y., Li, J., Zhang, Y., Zhou, X., Xu, L., Chen, L., Bi, G., Hu, X., Xu, F., & Wang, L. (2015). Processing of visually evoked innate fear by a non-canonical thalamic pathway [Publisher: Nature Publishing Group]. *Nature Communications*, *6*(1), 6756. <https://doi.org/10.1038/ncomms7756>
- Wekselblatt, J. B., Flister, E. D., Piscopo, D. M., & Niell, C. M. (2016). Large-scale imaging of cortical dynamics during sensory perception and behavior. *Journal of Neurophysiology*, *115*(6), 2852–2866. <https://doi.org/10.1152/jn.01056.2015>
- Weliky, M., Bosking, W. H., & Fitzpatrick, D. (1996). A systematic map of direction preference in primary visual cortex. *Nature*, *379*(6567), 725–728. <https://doi.org/10.1038/379725a0>
- Weng, X., & Kitani, K. (2019, August). Monocular 3D Object Detection with Pseudo-LiDAR Point Cloud [arXiv:1903.09847 [cs]]. <https://doi.org/10.48550/arXiv.1903.09847>
- Wheatstone. (1839). Ueber das Sehen mit zwei Augen und das Stereoskop. *Annalen der Physik*, *123*(8), 625–627. <https://doi.org/10.1002/andp.18391230813>
- Wheatstone, C. (1838). Contributions to the Physiology of Vision. Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision.

- Philosophical Transactions of the Royal Society of London*, 128(128), 371–394.
<https://doi.org/10.1098/rstl.1838.0019>
- Wheatstone, C. (1842). Beiträge zur Physiologie des Gesichtssinnes. *Annalen der Physik und Chemie*, 131(S1), 1–48. <https://doi.org/10.1002/andp.18421310102>
- Wheatstone, C. (1949). Binocular Depth Perception. In *Readings in general psychology* (pp. 9–20). Prentice-Hall, Inc. <https://doi.org/10.1037/11352-003>
- Wilson, B. J., Decker, K. E., & Roorda, A. (2002). Monochromatic aberrations provide an odd-error cue to focus direction [Publisher: Optica Publishing Group]. *JOSA A*, 19(5), 833–839. <https://doi.org/10.1364/JOSAA.19.000833>
- Wilson, J. J., Alexandre, N., Trentin, C., & Tripodi, M. (2018). Three-Dimensional Representation of Motor Space in the Mouse Superior Colliculus. *Current Biology*, 28(11), 1744–1755.e12. <https://doi.org/10.1016/j.cub.2018.04.021>
- Wilson, M. A., & McNaughton, B. L. (1993). Dynamics of the Hippocampal Ensemble Code for Space [Publisher: American Association for the Advancement of Science]. *Science*, 261(5124), 1055–1058. <https://doi.org/10.1126/science.8351520>
- Wundt, W. M. (1862). *Beiträge zur theorie der sinneswahrnehmung* [Google-Books-ID: MG8PAAAAQAAJ]. C. F. Winter.
- Wurtz, R. H., & Goldberg, M. E. (1972). Activity of superior colliculus in behaving monkey. IV. Effects of lesions on eye movements. [Publisher: American Physiological Society]. *Journal of Neurophysiology*, 35(4), 587–596. <https://doi.org/10.1152/jn.1972.35.4.587>
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research*, 48(20), 2070–2089. <https://doi.org/10.1016/j.visres.2008.03.021>
- Wurtz, R. H., & Kandel, E. R. (2000). Central visual pathways. In *Principles of Neural Science* (4th ed., pp. 523–545). McGraw-Hill Education.
- Xue, J., Ramoa, A., Carney, T., & Freeman, R. (1987). Binocular interaction in the dorsal lateral geniculate nucleus of the cat. *Experimental Brain Research*, 68(2). <https://doi.org/10.1007/BF00248796>
- Xue, M., Atallah, B. V., & Scanziani, M. (2014). Equalizing excitationinhibition ratios across visual cortical neurons [Publisher: Nature Publishing Group]. *Nature*, 511(7511), 596–600. <https://doi.org/10.1038/nature13321>
- Yerkes, R. M. (1904). Space perception of tortoises. *Journal of Comparative Neurology and Psychology*, 14(1), 17–26. <https://doi.org/10.1002/cne.920140103>
- Yildirim, M., Sugihara, H., So, P. T. C., & Sur, M. (2019). Functional imaging of visual cortical layers and subplate in awake mice with optimized three-photon microscopy [Publisher: Nature Publishing Group]. *Nature Communications*, 10(1), 177. <https://doi.org/10.1038/s41467-018-08179-6>
- Yilmaz, M., & Meister, M. (2013). Rapid innate defensive responses of mice to looming visual stimuli. *Current biology : CB*, 23(20), 10.1016/j.cub.2013.08.015. <https://doi.org/10.1016/j.cub.2013.08.015>

- Ylimäki, M., Heikkilä, J., & Kannala, J. (2018). Accurate 3-D Reconstruction with RGB-D Cameras using Depth Map Fusion and Pose Refinement [ISSN: 1051-4651]. *2018 24th International Conference on Pattern Recognition (ICPR)*, 1977–1982. <https://doi.org/10.1109/ICPR.2018.8545508>
- Yoonessi, A., & Baker, C. L., Jr. (2011). Contribution of motion parallax to segmentation and depth perception. *Journal of Vision*, *11*(9), 13. <https://doi.org/10.1167/11.9.13>
- Yoonessi, A., & Baker, C. L., Jr. (2013). Depth perception from dynamic occlusion in motion parallax: Roles of expansion-compression versus accretion-deletion. *Journal of Vision*, *13*(12), 10. <https://doi.org/10.1167/13.12.10>
- You, Y., Wang, Y., Chao, W.-L., Garg, D., Pleiss, G., Hariharan, B., Campbell, M., & Weinberger, K. Q. (2020, February). Pseudo-LiDAR++: Accurate Depth for 3D Object Detection in Autonomous Driving [arXiv:1906.06310 [cs]]. <https://doi.org/10.48550/arXiv.1906.06310>
- Yu, F., & Gallup, D. (2014). 3D Reconstruction from Accidental Motion. *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 3986–3993. <https://doi.org/10.1109/CVPR.2014.509>
- Yu, Y., Stirman, J. N., Dorsett, C. R., & Smith, S. L. (2022). Selective representations of texture and motion in mouse higher visual areas. *Current Biology*, *32*(13), 2810–2820.e5. <https://doi.org/10.1016/j.cub.2022.04.091>
- Zeater, N., Cheong, S. K., Solomon, S. G., Dreher, B., & Martin, P. R. (2015). Binocular Visual Responses in the Primate Lateral Geniculate Nucleus [Publisher: Elsevier]. *Current Biology*, *25*(24), 3190–3195. <https://doi.org/10.1016/j.cub.2015.10.033>
- Zhang, A., Jin, L., Yao, S., Matsuyama, M., van Velthoven, C. T. J., Sullivan, H. A., Sun, N., Kellis, M., Tasic, B., Wickersham, I., & Chen, X. (2024). Rabies virus-based barcoded neuroanatomy resolved by single-cell RNA and in situ sequencing. *eLife*, *12*, RP87866. <https://doi.org/10.7554/eLife.87866>
- Zhang, J., Zhang, X., Hu, X., Wu, W., & Yang, Y. (2017). Organization of spatial frequency in cat striate cortex. *Neuroscience*, *362*, 95–103. <https://doi.org/10.1016/j.neuroscience.2017.08.021>
- Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G. J., Tsang, A., Tsegaye, G., Narayan, S., Obara, C. J., Lim, J.-X., Patel, R., Zhang, R., Ahrens, M. B., Turner, G. C., Wang, S. S.-H., Korff, W. L., . . . Looger, L. L. (2023). Fast and sensitive GCaMP calcium indicators for imaging neural populations [Publisher: Nature Publishing Group]. *Nature*, *615*(7954), 884–891. <https://doi.org/10.1038/s41586-023-05828-9>
- Zhang, Y., Kim, I.-J., Sanes, J. R., & Meister, M. (2012). The most numerous ganglion cell type of the mouse retina is a selective feature detector [Publisher: Proceedings of the National Academy of Sciences]. *Proceedings of the National Academy of Sciences*, *109*(36), E2391–E2398. <https://doi.org/10.1073/pnas.1211547109>

- Zhang, Y., & Schiller, P. H. (2008). The effect of overall stimulus velocity on motion parallax. *Visual Neuroscience*, 25(1), 3–15. <https://doi.org/10.1017/S0952523808080012>
- Zingg, B., Chou, X.-l., Zhang, Z.-g., Mesik, L., Liang, F., Tao, H. W., & Zhang, L. I. (2017). AAV-mediated anterograde transsynaptic tagging: Mapping corticocollicular input-defined neural pathways for defense behaviors [Publisher: Elsevier]. *Neuron*, 93(1), 33–47.
- Zipser, D., & Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons [Publisher: Nature Publishing Group]. *Nature*, 331(6158), 679–684. <https://doi.org/10.1038/331679a0>
- Zmarz, P., & Keller, G. B. (2016). Mismatch Receptive Fields in Mouse Visual Cortex. *Neuron*, 92(4), 766–772. <https://doi.org/10.1016/j.neuron.2016.09.057>
- Znamenskiy, P., Kim, M.-H., Muir, D. R., Iacaruso, M. F., Hofer, S. B., & Mrsic-Flogel, T. D. (2024). Functional specificity of recurrent inhibition in visual cortex [Publisher: Elsevier]. *Neuron*, 112(6), 991–1000.e8. <https://doi.org/10.1016/j.neuron.2023.12.013>