BROADENING PERCEPTION: NATURE INSPIRED BROADBAND ORIENTATION STIMULI IMPROVE NEURONAL REPRESENTATIONS AND VISUAL PERCEPTION IN MICE

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades einer Doktorin der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Elisabeta Balla, M.Sc.

aus Elbasan, Albanien

Berichter: Universitätsprofessor Dr. Björn Kampa
Universitätsprofessor Dr. Simon Musall

Tag der mündlichen Prüfung: 03.07.2025

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek verfügbar.

M.Sc. Elisabeta Balla: *Broadening perception: Nature inspired broadband orientation stimuli improve neuronal representations and visual perception in mice*, © 2025

ABSTRACT

The visual system has evolved to optimize the representation of natural visual signals and their statistics. Several aspects of natural visual feature compositions have been well explored and found to have a direct representation in the visual cortex of mammals. For example, neurons in the primary visual cortex (V1) are tuned to different edge orientations or to different spatial frequencies. An important aspect of natural scene statistics that remains understudied is how the broadness of distribution, in which these features are commonly found in nature, impacts cortical representation and perception of natural visual scenes.

To address how feature broadness affects cortical responses, we generated random phase texture stimuli with parametrically-defined bandwidths of orientations or spatial frequencies and presented them to awake mice. Neuronal responses of their primary visual cortex (V1) were recorded with the help of a 2-photon microscope. We saw that broadening the orientation bandwidth of the stimulus strengthened neuronal responses and neuronal recruitment, whereas spatial frequency bandwidth broadening didn't have such an impact.

To parse whether this effect was merely due to diversely orientation-tuned V1 neurons being recruited with more orientations being present in the stimulus, we built a Gabor filter bank model simulating V1 population responses and compared it under different regimes to real mouse V1 responses. Orientation tuning alone could not explain the broadband orientation response enrichment effect. Adding an orientation tuned divisive suppression enabled the model responses to match real V1 responses.

Stimuli that extend uniformly beyond the preferred receptive field (RF) patch of a neuron can have a suppressive effect in V1 neurons. This is thought to have something to do with the predictability of the stimulus in the center region based on the similarity to the surround region. We looked at the stimulus spatial auto-correlation and it did indeed decrease by broadening the orientation bandwidth of the stimulus, meaning the center became less predictable based on the surround area information.

The impact of surround suppression was then experimentally confirmed by additional 2-photon imaging of center RF focused stimuli versus full screen stimuli. A linear mixed effect model (LME) with multiple regressors helped us parse the contribution of

orientation tuning, feedforward input, and surround suppression. While all these components had an impact on the broadband orientation responses' enrichment, surround suppression was the key contributor.

Moreover, looking at the global wide-field responses further cemented the crucial role of V1 in broadband orientation representation and exposed a visible lateral leaning gradient of higher visual cortex areas involved. Additionally, looking at the spiking activity of neurons at different V1 depths via Neuropixels probes showed that this response enrichment is widespread across layers. Interestingly, deeper cortical areas stood out with their strong preferences for a stimulus that is rich in both orientation and spatial frequency bandwidths, hence most highly unpredictable.

To address how feature broadness affects visual perception, we looked at the impact that richer orientation bandwidth stimuli had in the separability of neural responses to two different central spatial frequencies. Additionally, we looked at impact that richer spatial frequency bandwidth stimuli had in the separability of neural responses to two different central orientations. Once again, orientation bandwidth enrichment stands out in the impact it has at the neuronal level, whereby it increases neuronal separability.

Finally, integrating the separability of such stimuli as a discrimination task in a mouse touchscreen chamber in the form of a 2-AFC (Two Alternative Forced Choice), revealed that this increased separability also impacts the perceptual level. Because the mice also get better at discriminating different central spatial frequencies upon orientation bandwidth enrichment.

These conclusions support a predictability and context-dependent surround modulation of visual information at the cortical level. Scenes that are highly unpredictable and rich in orientation features drive a reduction in surround inhibition, which allows stronger recruitment of V1 neurons, increased neuronal response strength, a better representation of the visual information, and improved visual perception.

ZUSAMMENFASSUNG

Das visuelle System hat sich dazu entwickelt, die Darstellung natürlicher visueller Signale und ihrer Statistiken zu optimieren. Mehrere Aspekte der Zusammensetzung natürlicher visueller Merkmale sind gut erforscht und haben eine direkte Darstellung in der Sehrinde von Säugetieren. So sind beispielsweise Neuronen im primären visuellen Kortex (V1) auf unterschiedliche Orientierungen oder auf verschiedene räumliche Frequenzen abgestimmt. Ein wichtiger Aspekt der Statistik natürlicher Szenen, der noch nicht ausreichend erforscht ist, ist die Frage, wie sich die Bandbreite der Verteilung dieser Merkmale, die in der Natur häufig vorkommen, auf die kortikale Repräsentation und die Wahrnehmung natürlicher visueller Szenen auswirkt.

Um zu untersuchen, wie sich die Bandbreite der Merkmale auf die kortikalen Reaktionen auswirkt, wurden zufällige Phasen-Texturreize mit parametrisch definierten Bandbreiten von Orientierungen oder räumliche Frequenzen erzeugt und wachen Mäusen dargeboten. Die neuronalen Reaktionen ihres primären visuellen Cortex (V1) wurden mit Hilfe eines 2-photonen-Mikroskops aufgezeichnet. Es zeigte sich, dass eine Erweiterung der Orientierungsbandbreite des Reizes die neuronalen Reaktionen und die Rekrutierung von Neuronen verstärkte, während eine Erweiterung der räumliche Frequenz Bandbreite keine solche Wirkung hatte.

Um herauszufinden, ob dieser Effekt lediglich darauf zurückzuführen ist, dass unterschiedlich orientierte V1-Neuronen rekrutiert werden, wenn mehr Orientierungen im Stimulus vorhanden sind, haben wir ein Gabor-Filterbankmodell erstellt, das die Reaktionen der V1-Population simuliert, und es in verschiedenen Regimen mit den realen V1-Reaktionen der Maus verglichen. Das rekrutieren von mehr Neuronen allein konnte die Anreicherung von Orientierungsreaktionen nicht erklären. Durch Hinzufügen einer auf die Orientierung abgestimmten divisive suppression konnten die Modellantworten mit den realen V1-Antworten übereinstimmen.

Stimuli, die gleichmäßig über das bevorzugte rezeptive Feld (RF) eines Neurons hinausgehen, können bei V1-Neuronen eine unterdrückende Wirkung haben. Es wird vermutet, dass dies mit der Vorhersagbarkeit des Reizes in der zentralen Region auf

der Grundlage der Ähnlichkeit mit der umgebenden Region zusammenhängt. Das Untersuchen der räumlichen Autokorrelation des Reizes zeigte in der Tat, dass diese abnahm, wenn die Orientierungsbandbreite des Reizes erweitert wurde, wodurch das Zentrum auf der Grundlage der Informationen aus dem Umgebungsbereich weniger vorhersehbar wurde.

Die Auswirkung der Surround-Suppression wurde dann experimentell durch zusätzliches 2-photonen-Imaging von fokussierten RF-Stimuli in der Mitte im Vergleich zu Vollbild-Stimuli bestätigt. Ein lineares gemischtes Modell (LME) mit mehreren Regressoren ermöglichte es, den Beitrag der Orientierungsabstimmung, des Feedforward-Inputs und der Surround-Suppression zu analysieren. Während alle diese Komponenten einen Einfluss auf die Anreicherung der breitbandigen Orientierungsreaktionen hatten, war die Surround-Suppression der Hauptfaktor.

Darüber hinaus untermauerte die Betrachtung der globalen Weitfeldantworten die entscheidende Rolle von V1 bei der breitbandigen Orientierungsrepräsentation und zeigte einen sichtbaren lateralen Neigungsgradienten der beteiligten Höheres Visuelles Areal (HVA). Die Untersuchung der Spike-Aktivität von Neuronen in verschiedenen Tiefen von V1 mit Hilfe von Neuropixels-Sonden zeigte außerdem, dass diese Anreicherung von Antworten über alle Schichten hinweg weit verbreitet ist. Interessanterweise zeichneten sich tiefere kortikale Areale durch ihre starke Präferenz für einen Stimulus aus, der sowohl in Bezug auf die Orientierungs- als auch auf die räumliche Frequenz Bandbreite reichhaltig und damit höchst unvorhersehbar ist.

Um herauszufinden, wie sich die Bandbreite der Merkmale auf die visuelle Wahrnehmung auswirkt, untersuchte ich, wie sich eine Vergrößerung der Orientierungsbandbreite auf die Trennbarkeit der neuronalen Reaktionen auf zwei unterschiedliche zentrale Raumfrequenzen auswirkt. Umgekehrt untersuchte ich, wie sich eine Vergrößerung der Bandbreite an räumliche Frequenzen auf die Trennbarkeit neuronaler Reaktionen auf zwei verschiedene zentrale Orientierungen auswirkt. Auch hier fällt die Anreicherung der Orientierungsbandbreite auf als Faktor der die Trennbarkeit auf neuronaler Ebene erhöht

TABLE OF CONTENTS

ABSTRACT **3**

ZUSAMMENFASSUNG **5**

TABLE OF CONTENTS **7**

LIST OF ABBREVIATIONS **9**

LIST OF FIGURES **10**

LIST OF TABLES **10**

1. INTRODUCTION **11**

PROLOGUE 11
1.1 BORN TO SEE: ARE WE PRE-PROGRAMMED FOR NATURE? 12
1.1.1 As we learn about nature, we learn about our visual system. *14*
1.2 COMPLEX STIMULI REVEAL NON-LINEAR BEHAVIORS IN THE VISUAL CORTEX NEURONS. 18
1.2.1 Complex stimuli improve stimulus specific responses in the cortex. *20*
1.2.2 Higher-order patterns and integration non-linearities translate across species. *21*
1.2.3 Higher-order patterns and integration nonlinearities in mice. *26*
1.3 STUDYING NATURAL VISION WITH NATURAL IMAGES 29
1.3.1 Studying natural vision with higher order textures. *31*
1.3.2 Studying natural vision with complex and composite textures. *34*
1.3.3 Studying natural vision with visual noise. *36*
1.4 WHAT IS STILL MISSING IN OUR UNDERSTANDING OF NATURAL VISION? 39
1.4.1 Population level processing and remaining gaps. *39*
1.4.2 The concrete input-output hierarchy: subcortex, layers and areas. *40*
1.4.3 A link to perception and behavior. *41*
1.5 HOW DOES THIS STUDY ENRICH THE FIELD? 42
1.5.1 Specific aims. *43*

2. RESULTS **45**

2.1 BROADENING THE ORIENTATION BANDWIDTH OF A VISUAL STIMULUS TO MATCH NATURE STATISTICS LEADS TO AN INCREASED RECRUITMENT AND RESPONSE AMPLITUDE IN MOUSE V1 NEURONS. 45
2.2 BROADENING THE SPATIAL FREQUENCY OF THE STIMULUS DOES NOT HAVE A STRONG IMPACT IN THE RESPONSES OF MOUSE V1 LAYER 2/3. 50
2.3 A GABOR FILTER MODEL THAT INCLUDES SURROUND SUPPRESSION CAN ACCURATELY PREDICT RESPONSES TO BROADBAND ORIENTATIONS. 55
2.4 SURROUND MODULATION PLAYS A KEY ROLE IN ACCURATELY PREDICTING HOW INDIVIDUAL NEURONS RESPOND TO BROADBAND STIMULI. 59
2.5 BROADBAND VISUAL STIMULI LEAD TO INCREASED NEURAL RESPONSES IN BOTH CORTICAL AND SUBCORTICAL AREAS. 66
2.6 BROADER ORIENTATION BANDWIDTH LEADS TO AN ENHANCED ABILITY TO DISCRIMINATE BETWEEN VISUAL STIMULI IN THE NEURAL NETWORK. 70
2.7 BROADER ORIENTATION BANDWIDTH LEADS TO AN ENHANCED ABILITY TO DISCRIMINATE BETWEEN VISUAL STIMULI IN THE MOUSE. 74

3. DISCUSSION **77**

3.1 THE REPRESENTATION OF ORIENTATION BANDWIDTH ENRICHMENT. 78

3.1.1 Mouse V1 network dynamically adjusts its orientation tuning selectivity in response to higher-order visual features. *78*
3.1.2 Surround suppression regime depends on the spatial predictability. *80*
3.1.3 Broadband orientation stimuli drive strong responses in several lateral HVAs, but the strongest effect is in the granular layers of V1. ..... *82*
3.1.4 How far can these effects be generalized? *84*
3.1.5 Broadband orientation stimuli improve visual perception. *85*
3.2 THE REPRESENTATION OF SPATIAL FREQUENCY BANDWIDTH ENRICHMENT. 85
3.2.1 Mouse V1 maintains stable activity despite richer spatial frequencies. *86*
3.2.2 Broadband spatial frequency stimuli have a subtle effect in V1. ..... *87*
3.2.3 Broadband spatial frequency show no impact on visual perception. *88*
3.3 GENERAL CONCLUSION AND OUTLOOK. 89

4. METHODS 93

4.1 EXPERIMENTAL ANIMALS AND PREPARATION. 93
4.1.1 Behavioral monitoring. *93*
4.1.2 Surgical procedures. *94*
4.1.3 Intrinsic imaging during surgery *95*
4.2 DATA COLLECTION APPROACHES 96
4.2.1 2-photon imaging. *96*
4.2.2 Widefield imaging. *97*
4.2.3 Electrophysiological recordings. *98*
4.3 VISUAL STIMULATION. 99
4.4 TOUCHSCREEN CHAMBER EXPERIMENTS. 101
4.4.1 The orientation discrimination task. *101*
4.4.2 The spatial frequency discrimination task. *102*
4.5 DATA ANALYSIS AND STATISTICS. 103
4.5.1 Widefield. *103*
4.5.2 Electrophysiology. *103*
4.5.3 2-photon imaging. *103*
4.5.4 Statistics. *104*
4.6 VISUALIZATION. 108
4.7 STATISTICAL ANALYSIS OF NATURAL IMAGES AND MOTION CLOUD STIMULI. 108
4.7.1 Orientation distribution analysis. *108*
4.7.2 Spatial frequency spectrum analysis. *108*
4.7.3 Stimulus predictability analysis. *109*
4.7.4 Stimulus separability. *109*
4.7.5 Coefficient magnitude statistics. *110*
4.8 MODEL FOR NEURAL RESPONSES OF V1 NEURONS. 110

DATA AND CODE AVAILABILITY 115

REFERENCES 116

APENDIX 134

ACKNOWLEDGEMENTS 141

DECLARATION OF AUTHORSHIP 143

LIST OF ABBREVIATIONS

V1	- primary visual cortex
HVA/HVAs	- higher visual areas
LM	- lateromedial area
LGN	- lateral geniculate nucleus
SC	- superior colliculus
SF	- spatial frequency
LME	- linear mixed effects model
cpd	- cycles per degree
PV	- parvalbumin interneuron
VIP	- vasoactive intestinal peptide interneuron
SOM	- somatostatin interneuron
L1	- layer 1 of the visual cortex
L2/3	- layer 2/3 of the visual cortex
L4	- layer 4 of the visual cortex
L5	- layer 5 of the visual cortex
L6	- layer 6 of the visual cortex
AU	- arbitrary units
ΔF/F	- delta Fluorescence over Fluorescence
FOV	- field of view
FFT	- Fast Fourier Transform
GCaMP6s	- genetically encoded calcium indicator (version 6s)
NA	- numerical aperture
ORI	- broadband orientation stimulus
SF	- broadband spatial frequency stimulus
PCA	- Principal Component Analysis
PS analysis	- Portilla, Simoncelli analysis
RF	- Receptive Field
s.e.m.	- standard error of the mean
TI: Sapphire	- Titanium-Sapphire (laser type)

LIST OF FIGURES

Figure 1. Nature images contain a lot of edges and these edges often have contextual co-dependence. 16
Figure 2. Conceptual illustrations of surround suppression and proposed cell circuitry. 19
Figure 3. Mouse visual areas and information input-output hierarchy. 28
Figure 4. Example stimulation protocols using square wave grating stimuli versus natural movies. 30
Figure 5. Example pipeline for generating spectrally matched noise stimuli. 33
Figure 6. Example Motion Cloud movie rendering. 38
Figure 7. Broad orientation bandwidth stimuli drive strong and ample responses in the mouse V1. 46
Figure 8. The orientation bandwidth effect is consistent across animals. 48
Figure 9. V1 neurons maintain the same overall response despite spatial frequency bandwidth richness. 51
Figure 10. Spatial frequency bandwidth enrichment doesn't impact responses across mice. 53
Figure 11. V1 tuning is comparable across our tested range or orientations and spatial frequencies. 54
Figure 12. Gabor filter model links orientation tuning and surround suppression to broadband responses. 57
Figure 13. Broadband spatial frequency response modulation is not linked to the orientation tuning. 58
Figure 14. Broadband visual stimuli modulate neural responses by reducing surround suppression. 60
Figure 15. Pupil movements do not influence neural responses to stimuli presented at the RF center. 61
Figure 16. Neural response amplitudes were higher in the full-field mixed broadband condition. 62
Figure 17. Broadband orientation stimuli increase neural responses during different behavioral states. 64
Figure 18. Broadband visual stimuli increase sensory responses in cortical and subcortical neurons. 67
Figure 19. Expanding orientation bandwidth enhances stimulus discriminability in V1 neurons. 71
Figure 20. Bandwidth selectivity varies depending on the central spatial frequency or orientation. 72
Figure 21. Increasing orientation bandwidth leads to improved visual perception. 75
Figure 22. Schematic of the steps of the surgical preparation for imaging and headbar placement. 95
Figure 23. Example AUC for a strong and a weak central SF discriminator neuron. 107

Figure S1. Gabor filter bank model to represent mouse V1 responses. 134
Figure S2. Fitting Gabo model responses to narrow and broad orientations to the mouse V1 responses. 135
Figure S3. Stimulus separability estimated by coefficient magnitude statistics. 136
Figure S4. Broadband responses are different in the superficial and the deeper V1 neurons. 137
Figure S5. Neurons in the optical SC preferentially respond to the mixed broadband stimulus condition. 138
Figure S6. Retinotopic mapping of cortical visual areas and histology of the Neuropixels probe depth. 139
Figure S7. Learning curves of central spatial frequency discrimination in our mice. 140

LIST OF TABLES

Table 1. Specific Parameters for the orientation and spatial frequency band expanding protocols. 100

1. INTRODUCTION

Prologue

Imagine chefs in a restaurant. You see them chopping, smelling, sharpening knives, tasting, stirring etc. They pick some of the fresh ingredients presented to them, and ignore others. They rest, but then, more and more orders come in, so they start doing all these stereotypical actions at a higher pace and random-seeming grouping. To an observer without a broader understanding of a restaurant, this is all very bizarre activity. Being aware that the chefs operate within a larger function to provide customers delicious meals on demand, would be immensely helpful for the observer. The need for such awareness generalizes globally in quests to study or understand things. Trying to make sense of a combination of patterns without their context is in general very challenging and potentially giving rise to wrong assumptions or resorting to the “everything-bin” of vague categorization, like for e.g. “is it art?” or “probably just a religious ritual then”. Neuroscientists seek to understand the brain in depth, and have already managed to break it down to its fundamental functional units. But that, per se, didn’t simplify things; quite the opposite. Looking at neurons in the brain, you see them spiking, releasing chemicals, refining connections, pruning synapses, integrating inputs etc. They pick some of the stimuli presented to them, and ignore others. They rest, but then, the stimuli get more intense, so they start doing all these stereotypical firings at higher rates and random-seeming grouping. Touché, the unjustness of paralleling the brain with merely the hospitality sector of the economy while it is in fact an entity responsible for everything from body healthcare, to mind entertainment and governing hasn’t escaped my attention. But the fact is, typically in this field of study, one could realistically only focus on a narrow niche of research interest. Mine is the processing of the visual sensory information presented to the brain. Thus, it is important for me and my peers to understand who the customer is, what they want, what recipes the neuron chefs pick to cater to the customer’s demands and how they execute those. Also, the recipes are written in math, and in lack of a Rosetta stone for directly translating the neurons’ biological behaviors into mathematical operations, we will have to build computational analysis tools and models to do so along the way.

1.1 Born to see: are we pre-programmed for nature?

We are staying with the chef analogy for a little bit as we build two hypothetical scenarios. Exhibit A: we have a Szechuan cuisine chef, who has spent a lifetime manically perfecting an exclusive repertoire of two poetically spicy dishes. Their fingertips shriveled like Szechuan peppercorn, their breath smells like chili and they dream of the fire dragon cocooning their entire restaurant in a málà* cloud. Exhibit B: we have an equally skilled, general Asian cuisine restaurant owner/chef who is covering a wide spectrum of dishes and spices in their daily menu. Also, they happily tone down the spiciness level of their dishes to fit the untrained palette of their clientele. Experience has taught us to value more the first restaurant model, and that holds true in an unchanging environment. However, that model is also the most fragile as it is heavily dependent on stability of resources and consistent demand. The less specialized restaurant model is versatile, and probably better at improvising, integrating and generalizing dishes. However, a balance between both strategies is what usually makes a restaurant most successful. Keeping the authenticity level as high as possible, while accommodating the customer's (reasonable) demands, is the secret sauce. Similarly, a computation machine, such as the brain needs not only be highly specialized for perfection in particular tasks and representations, but also needs to be capable to generalize beyond statistical regularities while remaining plastic.

The main premise of today's vision neuroscience is that perception reflects the statistical structure of nature (Girshick et al., 2011; Simoncelli & Olshausen, 2001), shaping the brain much like a computational system which is optimized for a specific task. However, taken recent findings about processing of optical illusions, there have been calls to reevaluate this view and update in to a more utilitarian version thereof. Meaning, to consider perception as a product of empirical associations and biological utility, rather than a direct reflection of the physical world (Purves et al., 2015). Just as the general restaurant chef doesn't need to stay truthful to the original spiciness level of the dish to present it to their flavor minimalist customers, the utilitarian brain also doesn't need to be objective in representing nature, only capture what's useful to them.

---

* málà is a native Szechuan province seasoning consisting of Szechuan pepper corn and chili. Literally translated it means numbing (麻- má) and spicy (辣 – là).

And what is useful to them is what has shown useful to them. Meaning, the history of exposure to visual stimuli is what defines the perceptually relevant aspects of the input. The “pre-programmed for nature” and the “utilitarian brain” views don’t need to be mutually exclusive though. If we had evolved to only objectively represent nature, considering that nature statistics has remained relatively stable over evolutionary timescales, we should expect to find all truths about nature hardwired in the brain from birth as a statistically tuned computer. There are indeed a few instances of innate structures and biases found represented in the brain (which I will be describing in the following chapters). Yet, the system retains enough plasticity to refine itself based on actual experience. A lot of these refinements happen particularly during development. Furthermore, cross-cultural studies (McWhinnie et al., 1970; Thierry et al., 2009) have shown that experience and social contexts such as language can change susceptibility to certain colors or visual illusions. So, overall, the hardwiring is not the full story. Nature, and especially habitats or particular environments may vary from each other, and often offer habitat-specific non-expectancies, so it would make sense for the brain to mainly carry down in evolution the computations that are primed for universally stable regularities. Consistently, the visual system is biased towards natural environments and comes with built-in priors for efficiently processing the natural world, but it is not strictly limited to these priors and develops to calibrate itself in an experience dependent manner (Fiser et al., 2016; G. B. Keller & Mrsic-Flogel, 2018; Rao & Ballard, 1999; Simoncelli & Olshausen, 2001). To some extent, it can also can adapt beyond natural scenes e.g. the post-industrial revolution landscape.

Furthermore, our very own ability to generalize, recognize outliers, and construct internal models of the world hints at a balance between innate constraints and experience-driven learning. Being, not only a passive viewer of natural patterns, but a dynamic interpreter thereof is likely the key to this synergic balance. Finally, understanding what nature demands of the brain to represent, as we learn how the brain extracts, generalizes, and refines natural image statistics, will ultimately help us decipher the balance between stability and plasticity, shedding light on why on many levels, biological vision can outperform even the most advanced computational models (Susan, 2024).

1.1.1 As we learn about nature, we learn about our visual system.

Natural signals are known to be highly structured and contain within them a lot of repetitive or predictable information. This redundancy applies to many different types of signals, from sounds, climate, fractal structures and all the way to the evolution of human language motifs, whereby a very small amount of words or inflections is most abundantly used (Zipf, 1950). Natural images, making no exception to this, follow a power law distribution of spatial frequencies (Field, 1987), meaning there is more power contained in low frequencies than in high frequencies (1/frequency scaling). So, while small (fine) structures are more abundant, their individual contrast is tiny, and their power adds us slowly compared to the dominant contribution from large (coarse), high-contrast structures. Therefore, coarse structures dominate the energy in natural images. At the same time, while most of the image's power or variance is concentrated in slow, smooth transitions, those are also the parts of the image that contain the most redundant information. So, with nature being presented to us in such high resolution, we already meet the first challenge the visual system is faced with, the biological bottleneck to what information the system can process, relay and ecologically work with. It has been estimated that the retina receives about 10 million bits of information per second but only 10% of that can be transmitted by the optic nerve (Atick & Redlich, 1992). Furthermore, permanently processing that entire information would be slow and not robust enough for allowing perception, memory formation and other important neuronal phenomena to seamlessly take place (Olshausen & Field, 1996).

Thus, Barlow's first principle of efficient coding, which states that the brain first needs to remove redundant information, becomes very relevant in the first stages of sensory information processing (Barlow, 1961). Upon efficient coding being empirically introduced within the context of visual processing, it was formalized that early stages of this operation implement a form of spatial information decorrelation in order to reduce the amount of computational resources that need to be invested (Field, 1987). The reason this would be very efficient is because since nearby pixels are highly correlated, much of the variation in information is contained in the large-scale brightness variations rather than small details. Furthermore, it was proposed that the retina and LGN already remove redundancy and strong spatial correlations from natural images by performing a whitening transformation that enhances contrast at

different spatial scales and emphasizes edges rather than uniform areas (Atick & Redlich, 1992). In this model, the center-surround receptive fields of ganglion cells with their ON-center/OFF-surround and OFF-center/ON-surround responses act as high-pass filters which allow for the optimal level of whitening that is enough to amplify high frequencies without introducing too much high frequency noise in the system and keeping a good signal to noise ratio. It has later been shown (Rucci & Poletti, 2015) that gaze control through head or eye movements also plays an active role in these operations.

The primary visual cortex (V1) was also probed for this whitening transformation and it was shown that V1 simple cells behave like whitening filters, extracting features that minimize redundancy in natural images and allow focus on edge detection (Olshausen & Field, 1996). Additionally, using unsupervised learning, the authors of that same study showed that these V1 filters emerge naturally when optimizing for sparse coding. Gabor filters resemble oriented, localized wavelets of a defined scale and orientation tuning which makes them very well suited for the detection of oriented edges (Daugman, 1980). To some degree there is biological plausibility to the Gabor analogy for V1 neurons because receptive fields of neurons in V1 closely resemble Gabor filters in both structure and tuning properties (Hubel & Wiesel, 1959; Niell & Stryker, 2008). And since their responses are specific, (i.e. out a whole bank of Gabor filters with different orientations or spatial frequencies only a few of them respond to a given image patch) they mesh well with the sparsity of responses found in V1 neurons. It is, however, important to point out, that although the Gabor energy model captures well a lot of V1 neuron properties such as orientation-tuning and phase invariant responses, paralleling V1 to a homogenous bank of Gabor filters is a gross oversimplification. There is a considerable population of neurons that behave very differently from simple cells and have a diversity of receptive field shapes, combinations, preferences and perform non-linear integrations such as being tuned to specific feature conjunctions, higher order features etc. (Movshon et al., 1978; Skottun et al., 1991; Smyth et al., 2003).

Furthermore, looking at the hierarchical design of the visual cortex and the influence that higher visual areas have shown on modulation of V1 neuronal responses at particular contexts via gain control (Carandini et al., 1999; Sit et al., 2009) or pooling (Carandini & Heeger, 2012; Sakai & Tanaka, 2000), it is clear that there is more than

just whitening and edge detection happening in the first step of visual information processing and that non-linear operations happen in V1.

The understanding of whitening in the initial stages of visual processing was enriched furthermore, by looking into higher order dependencies and showing that structures such as edges or textures have conditional dependence i.e. the presence of an edge of a given orientation influences the probability of encountering another edge of similar orientation nearby, so it is easier to predict the continuation of a curvature based on one segment of it (Simoncelli & Olshausen, 2001). This means that edges found in our natural environment are usually not isolated. For example, a tree branch is more likely to have edges that run across it in the same direction (see Fig. 1) and a face is more likely to follow a smooth contour rather than sharp edges. On the one hand, edges repeat across the visual space like trees in a forest, hence building spatial frequencies. And on the other hand, they strongly overlap creating more textures or contours rather than sharp edges or lines. This highlights that perfect whitening is neither biologically achievable, nor desirable, because some correlations are actually useful for higher order texture inferences.

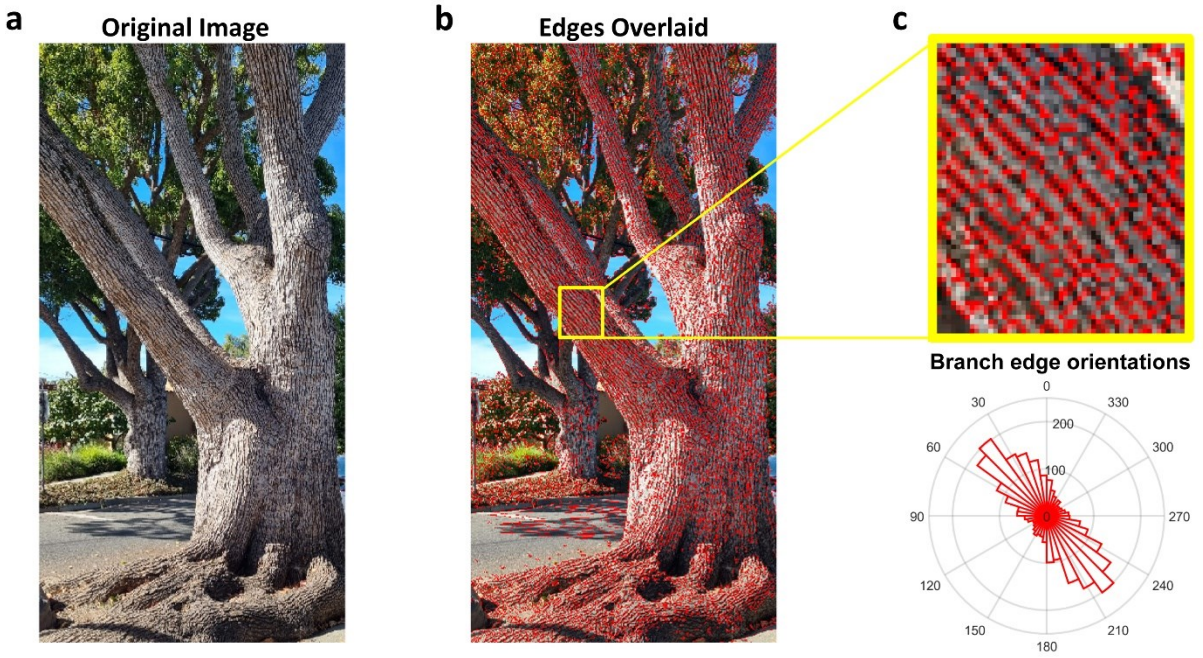


Figure 1. Nature images contain a lot of edges and these edges often have contextual co-dependence.

(**a**) Example image of a tree. (**b**) The edges of the image have been detected and overlaid on the original image as a red mask. (**c**) Top, example section of a branch from panel b (yellow square). Bottom, polar plot of orientation distributions in the selected image section. The edges in the brand tend to have similar orientations and the main direction that the brand grows towards has a large influence of these orientations.

Furthermore, looking at orientations, it has been shown that some of them are more predominantly found in nature: namely the horizontal and vertical edges frequently found in the horizon, tree edges and so on (Girshick et al., 2011). Neurons in V1 respond preferentially to a limited range of orientations (Hubel & Wiesel, 1959; Niell & Stryker, 2008). This tuning to particular orientations has been described extensively in several research models (Niell & Stryker, 2008; Priebe, 2016; Roth et al., 2012) with all their particularities. Generally, V1 shows a high level of orientation, direction, spatial frequency and speed tuning (Roth et al., 2012). Interestingly, the predominance of cardinal orientations finds its representation both in the population of cells that are tuned to these orientations in V1, but also in the perception of human subjects. Human perception has been shown to be biased towards cardinal orientations and tend to converge approximating orientations towards these dominating ones (Girshick et al., 2011). Furthermore, recent environment also seems to have an impact as increased exposure to a certain orientation can increase representation thereof in the orientation tuned population of V1 in mice (Kreile et al., 2011) rats (O'Hashi et al., 2007) and kittens (Sengpiel et al., 1999; Tanaka et al., 2007).

To summarize, we now know that there are high spatial correlations in natural images and that it is easy to predict information content based on neighboring regions, especially in smooth spaces where the coarse structures dominate. Furthermore, due to set physical properties such as the similar reflectance of surfaces, source of illumination etc., these correlations are scale invariant (Field, 1987; Stephens et al., 2013), meaning they will be found whether one looks up closely or on a zoomed-out perspective. We know that sharp luminance intensity changes that define boundaries (edges) and textures in natural images are highly structured rather than randomly occurring (Sigman et al., 2001) and edge orientations not only conditionally depend on each other but can often be perceived as combinations of common orientations such as for example horizontal and vertical edges which are also more predominantly present in our environment. Furthermore, natural scenes have a broad distribution of orientations (Girshick et al., 2011; Schwartz & Simoncelli, 2001). The retina and LGN (Dan et al., 1996) aided by active gaze control (Rucci & Poletti, 2015; Segal et al., 2015) perform some limited bandpass whitening of the visual input, equalizing the power within a certain range of spatial frequencies and temporal scales. V1 mainly sparsens responses and performs non-linear operations that are context relevant. All

these together, bring a lot of sense to why edge detection is such a fundamental operation in early vision as it is minimizing redundancy in natural inputs and mainly encoding for the most informative features.

It is important to note that all these statistical properties described in natural scenes differ strongly from those of artificial stimuli like white noise or gratings or bars, which mostly come with dominating single features and lack higher order correlations. This single feature dominance makes an artificial stimulus such as a sinewave grating drive neuronal responses in a stereotyped way (e.g. only activating the subset of neurons responding to the central orientation of the grating), whereas the complexity of a natural scene would activate the subset of neurons detecting specific features in it and would be more likely reveal further nonlinear integration properties of these neurons (for example, neurons that remain silent when the grating is shown could be active when that grating is part of a larger context or a coherent shape).

1.2 Complex stimuli reveal non-linear behaviors in the visual cortex neurons.

Throughout the years, many examples have been gathered that make a strong case for the claim that simplistic, artificial stimuli are lacking in describing visual processing of neurons. As it was initially observed that V1 neurons respond strongly to artificial stimuli like oriented bars or sinewave gratings, a lot of early models were built around these responses (Hubel & Wiesel, 1959; Niell & Stryker, 2008; Priebe, 2016). Simple cells have linear summation in responses when it comes to oriented gratings. Complex cells, however, don’t do this, and they are invariant to phase shifts in a stimulus. Hypercomplex cells show an even more interesting behavior: when stimulus extends beyond their preferred receptive field, they reduce their activity (Hubel & Wiesel, 1965).

On the one hand, we have neuronal activity reflecting integration of inputs over space and time to amplify responses to continuous or grouped elements (Kirchberger et al., 2021), on the other hand, there is surround suppression which is a reduction of responses to a uniform pattern when it extends beyond the classical receptive field of that neuron (Adesnik et al., 2012; A. J. Keller, Roth, et al., 2020). (See Fig. 2.)

However, placing a contrasting version of this pattern in the surround (e.g. an orthogonal orientation) can release the cell from surround suppression (Di Santo et al., 2022; A. J. Keller, Dipoppa, et al., 2020).

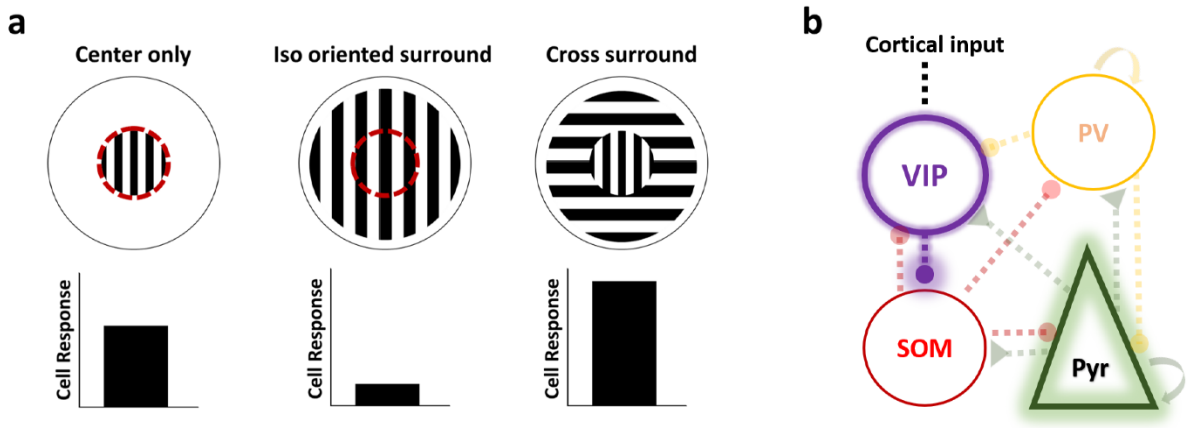


Figure 2. Conceptual illustrations of surround suppression and proposed cell circuitry.

(**a**) Conceptual illustrations of how the surround of the receptive field of a V1 neuron can modulate cell responses. While iso-oriented surround (middle) causes surround suppression, cross oriented surround (right) causes facilitation due to its potential pop-out effect. Red circle marks the receptive field of the neurons. (**b**) Schematic of the key players of the proposed surround suppression neuronal circuitry. The origin of the arrows shows agency and the ending shows upon which other cell type the inhibition is applied to. The pyramidal neuron (Pyr) is inhibited by the somatostatin interneuron (SOM), which in turn can be inhibited by the vasoactive intestinal peptide (VIP) interneuron. This disinhibits the pyramidal neuron and increases responses to salient stimuli.

Furthermore, it has also been shown that standard receptive field models fitted from grating responses fail to predict responses to dynamic natural scenes, while fitting on natural stimuli could predict time-varying V1 responses twice as accurately instead (David et al., 2004). This indicates that natural scene statistics engage V1's non-linear mechanisms differently from simple patterns. For example, responses to the edge of an objects can be enhanced relative to the background in what is known as figure ground segregation (Lamme, 1995). Here, V1 contributes to the formation of a pop-out perception of the figure by emphasizing a sparse figure-related code and removing responses to the background or to contour discontinuities. The involvement of attentional mechanisms (Roelfsema et al., 1998) and higher cortical areas that do a global integration of V1 local feature signals is required for such a high-level percept. But the versatile nature of these nonlinearities hints that V1 is potentially involved in interacting processes that regulate visual perception dynamically.

Additionally, operations like unexpected input driven recurrent enhancement (G. B. Keller & Mrsic-Flogel, 2018) or suppression (G. B. Keller et al., 2012) and the fact that

neurons display adaptive gain control which is the adjustment of responses based on recent environment such as exposure to a high a contrast stimulus (Albrecht et al., 1984; Cody & Tzounopoulos, 2022) or stimulus relevance (A. J. Keller et al., 2017), leaves room for the presence of a prediction error coding which would take place faster than the processing of the raw sensory input and is fundamentally a nonlinear transformation. Finally, looking at a population level, much more of this becomes apparent. Studies like (Stringer et al., 2019; Whiteway et al., 2019) suggest that population responses in visual cortex are high-dimensional and nonlinear.

1.2.1 Complex stimuli improve stimulus specific responses in the cortex.

A stimulus has been encoded efficiently when its most important features have been represented while the lowest amount of resources have been used. Natural scenes tend to drive neuronal responses in a very efficient way (Bányai et al., 2019). V1 neurons are tuned to specific orientations, spatial frequencies and other features which can be found in natural scenes, but natural scenes and naturalistic stimuli lead to a sparse activation of a subset of these neurons (Tang et al., 2018).

Beside conserving resources, another advantage of this spare activation is that the potential overlap in population responses to different stimuli decreases, improving stimulus specificity. It has been demonstrated that higher-order correlations in natural scenes induce a sparser code in mouse V1. Here, the information is encoded by the reliable activation of a small set of neurons which facilitates read-out (Froudarakis et al., 2014). Additionally, it has been shown that responses of neurons in primary visual cortex are sparse and more variable during presentation of natural scene stimuli, suggesting a reduction in redundancy and an efficient coding strategy (Vinje & Gallant, 2000; Yoshida & Ohki, 2020).

Furthermore, natural scenes have shown to improve the signal to noise ratio of V1 population response. This is somewhat intuitive as the evolutionary tuning of the brain is aligned with the nature of these inputs, thus the responses would have a higher fidelity to these features and less noise (Froudarakis et al., 2014).

An increased synchronicity of activity in neurons with overlapping receptive field properties has also been observed in natural scene processing (Papale et al., 2024). This enhances the signal to the perceptual borders they are all commonly tuned to. Natural scenes provide rich contextual information which V1 neurons potentially leverage for building an internalization of natural patterns (Berkes et al., 2011) and for encoding deviations from expected patterns in the local stimulus information. For example neuronal firing rates in V1 have been shown to be modulated by the contextual predictability of higher-order image features, aligning with predictive coding models (Uran et al., 2022). Natural scenes have even been proven helpful in improving the accuracy and the robustness of visual motion detector models (Clark et al., 2014; Yu et al., 2022).

Finally, natural scenes are behaviorally relevant. This means that the hierarchical visual information processing model of the brain needs to be very much considered also in context of what happens in the initial V1 representations. Feedback coming from higher visual areas and attentional systems should help to enhance and improve the representation of natural scenes in V1. It could be doing so by modulating receptive field tuning or the response gain. Experiments with awake, behaving macaques (Lazar et al., 2024) found that when paired with attention demanding tasks, natural scenes evoke more specific and stronger V1 responses, meaning that the responses are simultaneously modulated by attention, context and scene statistics.

1.2.2 Higher-order patterns and integration non-linearities translate across species.

An interesting question is how universal are the coding strategies for natural scenes across different species? Most of today's research in this segment comes from mammals with a visual cortex and different levels of complexities. Comparative studies in mice, rats, cats, monkeys and humans suggest that a set of general principles exist but species-specific differences need to be considered as well. In all these cases, the visual cortices have neurons tuned to oriented edges, indicating that edge detection is a fundamental operation for all these mammals.

In rodents, V1 neuron orientation selectivity tuning tends to be broad and organized in a salt and pepper pattern rather than in smooth maps, while in cats and primates,

orientation tuning is sharp and organized into defined orientation maps (Hubel & Wiesel, 1959; Kaschube, 2014; Ohki et al., 2005; Swindale, 1996) . Yet, despite this strict and seemingly edge detection specific organization, neurons in the cat V1 columns have been shown to exhibit stimulus independent functional heterogeneity (Martin & Schröder, 2013; Yen et al., 2007), which suggests a diverse, context-dependent coding strategy.

Another shared principle is decorrelated and sparse population activity in response to natural stimuli. When recording from a large population of neurons of the mouse V1 while presenting them natural movies and decorrelated (phase scrambled) stimuli, the natural scenes evoked a distinctly sparse population code (Froudarakis et al., 2014). Each image from the video would strongly activate a small subset of neurons. Most other neurons would remain inactive. When higher-order correlations were removed from the stimuli, the responses became less sparse or specific and more neurons responded to each stimulus in general. This is also reflected in results from studies in the monkey. For example, in an earlies study, increased sparseness and reduced pairwise correlation was observed in monkey V1 in response to natural input (Vinje & Gallant, 2000). This pattern has also been observed in several cat studies (Baddeley et al., 1997; Yen et al., 2007). Thus, sparsening of responses to natural scenes appears to be a general principle and likely finds agency in the presence of nonlinear inhibitory circuits (e.g. divisive normalization) that have been conserved through evolution thanks to being advantageous for encoding natural visual inputs.

Furthermore, different types of nonlinear interactions such as surround suppression, cross-orientation suppression, and adaptive gain control have been found present across all these species. Pronounced surround suppression has been found mainly in cat and monkey V1 (Cavanaugh et al., 2002; Jones et al., 2001; Sengpiel et al., 1997, 1999); although instances of that are also now found in mouse studies as well (A. J. Keller, Dipoppa, et al., 2020; Vaiceliunaite et al., 2013; Vangeneugden et al., 2019). Some form of contrast gain control has also been found present in all these animal models. Circuit motifs that implement efficient coding (e.g. lateral inhibition for decorrelation, input normalization for dynamic range adjustment) are thus part of the mammalian visual cortex template overall.

As for humans, evidence to make a case that human perception is tuned to natural statistics has been collecting in the recent years. For example, an internal reflection of the non-independence of orientations in natural scenes is made apparent in a study that showed that the human visual system is more sensitive to edges that form continuous contours (Field et al., 1993). Studies that point out to the fact that our perception is biased towards dominant orientations in our visual conditioning such as horizontal or vertical edges (Girshick et al., 2011), and our enhanced contour integration skills, highlight how prevalent statistics have shaped our perceptual strategies. Additionally, fMRI/ EEG studies (Kohler et al., 2016) show that viewing natural-like textures versus random noise evokes different activation patterns and pathways in the human visual cortex, which allows space for the idea that the human visual system is also optimized for natural input.

Certainly, due to ecological challenges and biological build up, there are also differences taking place. A very glaring difference is the level of visual acuity across the species mentioned:

Primates have a much higher visual acuity and thanks to the presence of a central focus point in the retina; the fovea, which means a lot of V1 neurons are specifically capable to capture fine details in natural scenes and have a high spatial and temporal acuity. Generally, their visual cortex shows a sharply hierarchical build with feature-specific pathways (V2, V4, IT) where by V1, appears to fall more into the Gabor-like tuning function with their columnar organization. However, neurons that do more complex operations such as border ownership for figure-ground segregation are also found.

Cats also have a columnar organization with strong orientation selectivity (yet neurons in these columns have been shown (Martin & Schröder, 2013) to exhibit stimulus independent functional heterogeneity) and a relatively high visual acuity and high spatial resolution. The study of cat V1 neurons has helped establish concepts such as surround suppression and complex/hypercomplex cells.

Mice, on the other hand, have their eyes in the sides of their heads, which gives a large field of view (~270°-320°), is mainly monocular and only has a small (~40°) binocular overlap (Poort & Meyer, 2021). They do not have a fovea and have a much

lower visual acuity. Instead, their V1 is most likely focused on recognizing large objects and detecting motion. Instead of a columnar organization, it has a salt and pepper spread of tuned cells which seem to be functionally linked without physical clustering vis-à-vis orientation tuning (Hubel & Wiesel, 1959; Kaschube, 2014; Ohki et al., 2005; Swindale, 1996). Also, the hierarchical organization of the different visual cortex areas is less well defined compared to their carnivore counterparts in primates and cats. Although some feature specificity has been found, such as the higher visual area (HVA) in the lateromedial visual cortex (LM) and the anterolateral visual cortex (AL), serving as entry points for, respectively, ventral and dorsal processing streams (Q. Wang et al., 2011), and LM also showing a higher selectivity for textures (Bolaños et al., 2024). There is, however, still much to find about the strength the hierarchy and feature pathways. Mice have a tiny body and brain size in comparison, so the size of the different brain areas presents a different distance scale that needs to be covered by both recurrent and feedback signals, affecting their latencies. Also, behavioral states such as running have a strong, positively correlated effect on V1 activity in mice but not in primates (Liska et al., 2024), possibly indicating different vision to action relation. Finally, being nocturnal animals, mice are also endowed with a different set of photoreceptor distributions and cone types which paints a different picture of statistically prevalent and environmentally relevant visual signals. Yet, when these signals reach cortex, the computation seems consistent: extracting informative features and reducing redundancy.

To summarize, across several species, regardless of brain size or habitat, similar and efficient coding strategies are employed for processing of natural signals. The redundancy of the information is reduced via bandpass filter early on in the visual path. Neurons employ nonlinear suppression to generate sparser codes; and populations collectively achieve an efficient representation, that allows for important features only to be extracted. There are potentially two parallel lines that define how these principles find expression: long term evolutionary exposure to environments that have similar structures such as edges or textures and developmental plasticity and refinement of neuronal tuning can be refined by recent visual experience.

1.2.2.1. Development of nature representations is ecologically defined.

Studying a fully developed animal in regards to how a certain representation is established in their brain, is like taking a historical snapshot of the aftermath of an event. Though, richly informative, it is a temporally challenged perspective on how this establishment happened. What was nature, what was nurture? Is nature and nurture the same for different species with different sets of evolutionary pressures? One would need to take snapshots at several stages of development to have clearer perspective on the establishment of these computations. Let's look at a comparison between two example mammals: mice and ferrets.

Mouse V1 neurons show both orientation selectivity and direction selectivity right from eye opening (Rochefort et al., 2011). This implies that their visual system can in theory handle basic visual features such as edge detection and motion. Considering their active, exploratory behavior, which includes a lot of rapid head movements, this innate preparedness could be advantageous in representing the quickly changing visual field. Contrary to the innate basic feature processing, mice need visual experience to develop higher-order visual processing such as surround modulation (Pecka et al., 2014). This suggests that, mice had to embrace a flexible, experience-driven way of establishing natural statistics representation, where they carry the starter package for immediate motion detection (orientation and direction selectivity) but then undergo a visual experience driven fine-tuning to unlock their more advanced visual functions based on the specific environmental challenges and experiences.

Ferrets, on the other hand, are born with orientation selectivity but need a few weeks of visual experience to develop direction selectivity (Roy et al., 2018). In principle, their neurons seem to be already primed for direction selectivity but need visual experience to give permission for this capability to be accessed. Contrary to this, their ability to execute higher-order visual processes like center-surround and cross-orientation suppression is already there from eye opening (Popović et al., 2018). It is interesting that ferrets come with a toolbox for complex feature processing, whether that has something to do with their predatory nature is probably still to be determined. But one could see how capabilities like contrast processing and focusing on the relevant parts of the visual field could be good training grounds for a hunting animal. However, ferrets also keep perfecting their higher order representations of nature statistics throughout

their life, as it has been shown that with time their ongoing spontaneous activity becomes more similar to previous responses to natural scenes, suggesting the network keeps internalizing nature with more and more visual experience thereof (Berkes et al., 2011).

Both animals improve their internalization of higher order natural statistics with visual experience, but the specificity and timelines differ. This emphasizes that the previously mentioned balance between innate mechanisms and experience-driven refinement in visual processing. Lifestyle, feeding niche, evolutionary pressures and habitat-specific needs seem to drive or at least influence what part is fixed and is flexible for refinement and generalization of environmental patterns. Mice, might need to navigate more variable and unpredictable micro-environments, and need to prioritize early motion processing for the dynamic aspect of their lifestyle, while employing environmental specific higher feature fine-tuning to cope with the variability aspect of their habitat. While, ferrets might have more stable environmental conditions or specialized hunting niches, thus have some higher-order visual processing mechanisms pre-configured from the get-go, allowing them to optimize certain tasks more quickly. Also, since they have a comparatively higher visual acuity, more refinement of the spatial structure of the visual scene and minimization of the redundant information might be required before fine-tuning motion perception. Food for thought: if such an acuity driven gradient exists, one could extrapolate that mammals with even higher visual acuity such as primates, would be in most need of such a mechanism.

1.2.3 Higher-order patterns and integration nonlinearities in mice.

The model animal of choice in this study is the mouse (Mus Musculus), so it is important to look at their specificities when it comes to the visual system. The mouse is a very common experimental animal due to ethical and resource considerations, ease of handle and population growth rate. Mice have also been very well characterized in terms of anatomy, morphology and cellular build up. There are a lot of tools publicly or commercially available for optimizing the study of the mouse brain, such as areal atlases, virus suites for controlled expression of fluorescent proteins that allow us to read out or modify the responses of particular neurons. The literature is

also fast-growing catching up to the analogous studies conducted in more complex mammals such as cats and primates. Being fairly simpler in comparison makes mice a bit easier to study but also comes with its own limitations. Simpler also means, carrying translational gaps. Therefore, it is important to look in depth at which of these nonlinearities actually apply to the mouse and which are expected to translate but have not quite been confirmed experimentally, and what can be generalized.

The most obvious differences to us are the body size and the nocturnal agenda of mice. Foraging in the dark requires for other modalities than vision to be prioritized, hence a less ambitious visual system develops. Furthermore, considering their eye size and lack of fovea, their visual acuity is not as good and not capable to resolve very high spatial frequencies or very fine structures. Also, they don't have an obvious structural organization of orientation tuned neurons such as the columnar primate organization but have a scattered salt and pepper spread thereof (Kaschube, 2014; Ohki et al., 2005).

The mouse cortex is made of highly specialized layers which have their own input-output hierarchy (see Fig. 3). Layer 2/3 (L2/3), performs computations and complex non-linearities that are more crucial in low-level statistics. Therefore, it is what most of the upcoming sections of the introduction are based on. It is, however, important to keep in consideration that a bigger mechanism and circuit is at play for realizing these computations. Layer 4 (L4) is the first recipient of thalamic signals, thus considered the feedforward information carrier. Some orientation tuning and surround suppression is found also in this layer, suggesting subcortical areas could be potential contributors to some of these modulation effects (Self et al., 2014). Deeper layers such as layer layer 5 (L5) and layer 6 (L6) neurons interact with areas further away. L6 neurons send feedback to the thalamus and modulate help modulate input gain. Whereas L5 neurons often have large dendritic arbors which allow them to integrate more broadly as they sample from superficial layers. These neurons also show the strong context dependent response modulation consistent with the anatomy that HVA feedback targets their apical dendrites which extend all the way up to layer 1 (L1) (Burkhalter et al., 2023). The cortical layers execute different aspects of visual information processing: input layers detect and relay information about basic features, L2/3 and L5 integrate horizontal and feedback signals to extract higher-order patterns.

Furthermore, having information being forwarded from L2/3 towards the HVAs and fed back to the thalamus by the L6 enables mouse V1 to not only encode simple stimulus properties but also to dynamically adjust and integrate information for a context relevant processing of the complex statistical structure of natural scenes.

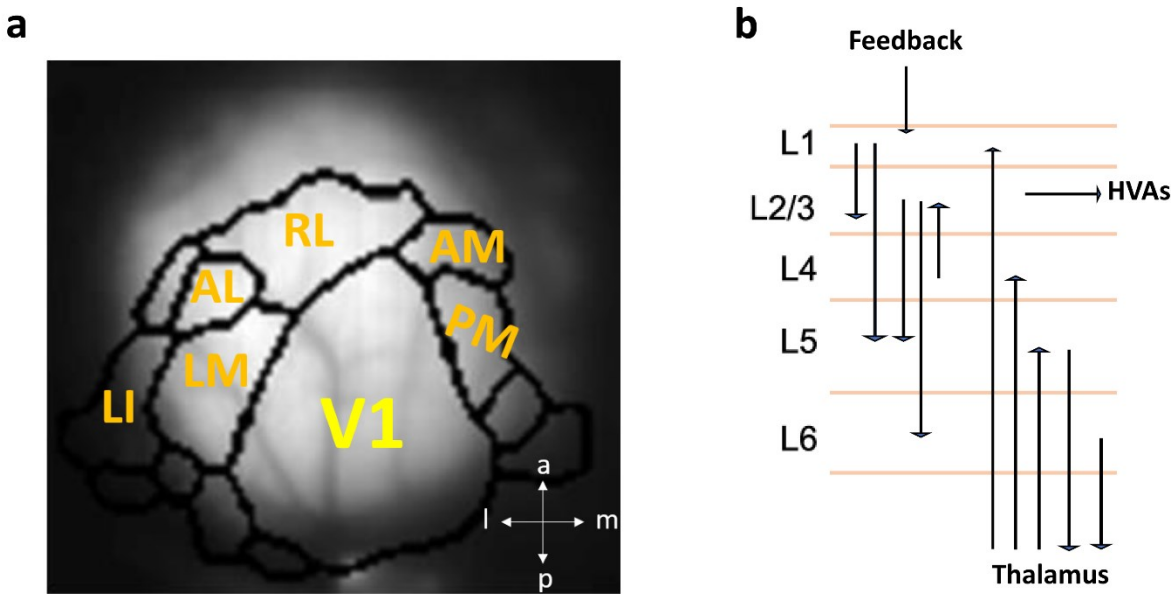


Figure 3. Mouse visual areas and information input-output hierarchy.

(**a**) Mouse visual cortical areas mapped on top of a wide-field view of the brain. Primary visual cortex (V1). Lateral areas: Rostro lateral (RL), Antero lateral (AL) Lateral Medial (LM), Lateral Intermediate (LI). Medial areas: Antero medial (AM), Posterior medial (PM). (**b**) Simplified scheme of the information flow through the V1 layers. Direction of the arrow shows the direction of the information flow.

Furthermore, mouse V1 neurons are selective for fundamental visual features such as orientations and spatial frequencies, they have the classic simple cell edge detector present along with complex-cell-like behaving neurons and interneurons (Bonin et al., 2011). And not only does one find non-linearities such as surround suppression in their V1, but also these linearities are expressed asymmetrically in the population of neurons. Some neurons are uniformly suppressed by the surround, some are end-stopped and some side-stopped (Samonds et al., 2017). It is not only responses to simple features that make mouse V1 neurons interesting. They also achieve complex figure-ground segregation (Kirchberger et al., 2021, 2023). An effect which seems to be strongly connected with the activity along a lateral path of HVAs with LM being the focal point. These contextual responses are potentially emerging from horizontal and HVA feedback signals and are engaging a disinhibitory circuit in V1 whereby figure regions preferentially activate vasoactive intestinal peptide (VIP) interneurons and

suppress (somatostatin) SOM interneurons, which releases pyramidal neurons from inhibition.

Concepts like surround modulation, pop-out or figure ground modulation, all appear as separate puzzle pieces of the big “contextual processing” picture. This involves integration of inputs over space and time to amplify responses to continuous or grouped elements (Kirchberger et al., 2021), along with reduction of responses to edges or uniform pattern when they extends beyond the classical receptive field of that neuron (Adesnik et al., 2012; A. J. Keller, Roth, et al., 2020), or the release from surround suppression upon a contrasting surround (A. J. Keller, Dipoppa, et al., 2020). Feedback from HVAs and the engagement of an inhibition/disinhibition (VIP - SOM - pyramidal neuron) circuit seems to be the recurring theme.

1.3 Studying natural vision with natural images.

Neuroscience is becoming more and more aware that the use of minimalistic artificial stimuli such as sinewave gratings of simple bars is very limited in what it can reveal about neuronal encoding of visual information. The feature distributions such stimuli can cover is limited and would require large-scale combinatorics to describe complex behavior of V1 neurons.

Intuitively one could ask why not just use natural scenes, animal videos or images to identify responses pertaining complex natural statistics? Certainly, the possibility of using natural images as experimental stimuli is a possibility which has (Antolík et al., 2016; Kayser, 2004; Stansbury et al., 2013) and is currently being explored in the field (See Fig. 4). Researchers have used these stimuli to build better V1 models or to look at species differences such as cortical organization driven activity patterns and feature selectivity. For example, spontaneous activity patterns are found to be closer in similarity to stimulus evoked responses in marmoset V1 and V2 than in the mouse V1, where these patterns are orthogonal (Matsui et al., 2024). Mouse V1 neurons also have been shown to exhibit heterogeneous orientation tuning across their receptive fields, indicating complex spatial feature selectivity emerges earlier in mice than in primates (Fu et al., 2024). Natural images have also been used to probe and predict

receptive fields of neurons, but for a large number of neurons the responses to a set of natural images deviated nonlinearly from the predicted responses which were based on the receptive fields that where determined by another set of natural images (Smyth et al., 2003).

Using natural images helped build an improved model, which considers shared feedforward input in the population for predicting neural responses to natural scenes, (Antolík et al., 2016). Furthermore, another mouse visual cortex model can accurately predict neuronal responses to novel natural stimuli, enabling better in silico experimentation (E. Y. Wang et al., 2025)

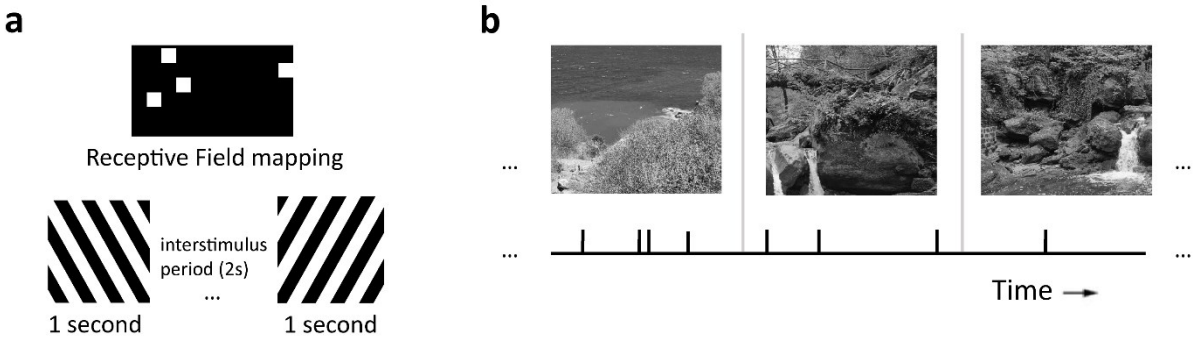


Figure 4. Example stimulation protocols using square wave grating stimuli versus natural movies.

(**a**) Example stimulation protocol used for 2-photon imaging, where the animal was shown full screen square wave gratings of different orientations. The neurons are mapped for their preferred area through a sparse noise stimulus (top) and then a randomized sequence of different square wave gratings of different orientations was shown. (**b**) Simplified illustration of a natural movie protocol. The three squares on top are snippets of the natural scenes from the movie and the vertical lines show schematic of a binned spike train.

However, the biggest pitfall of using natural images or movies is the lack of experimental control. Due to their large variabilities and co-occurrence of features in the same images, natural scenes are not easy to define in a parameter space (Felsen & Dan, 2005; Rust & Movshon, 2005). These countless parameters make it hard to pinpoint the specific features that drive a neuronal response or a sensory perception. It is easy to derive the exact feature that makes two gratings different in an experimental protocol, because it is usually the orientation or the spatial frequency, but that is not as straightforward with two different images of forests.

Another point to consider is how the variability to the overall network state could be affected through contrast and luminance changes in individual images. So, to encompass for the overall variability mitigation and large parameter space that needs

to be covered, a very large set of images of videos needs to be built and screened. This is very time consuming and hard to get and prolongs experimental duration. Furthermore, the analysis and interpretation of such large and complex data sets begs the use of advanced and robust linear-nonlinear models or deep neural networks to predict responses. The sparsity of V1 responses to natural images brings out the data-hungry aspect of these models (Benda et al., 2007), which is nor desirable nor sustainable.

1.3.1 Studying natural vision with higher order textures.

Due to all the limitations that come with the use of natural images libraries as stimuli in visual neuroscience, an alternative has been the use of stimuli that “look like nature”. As already established, natural scenes go beyond just a collection of edges and are composed of a battery of higher order statistical features such as textures, contours and shapes that are perceptually distinguishable from their background in a distinct figure-ground relationship. The more informative higher order statistics of natural images are, the more perceptually salient they are to human participants and can be better distinguished by them from unstructured noise (Tkačik et al., 2010).

So, in a quest to probe which particular aspects of these structures the neurons in the visual cortex are sensitive to, filters for particular features have been developed with the aim to use their output to synthetize desired artificial images in a controlled manner. For example, one can take a natural image and randomize their phase (this preserves the second order statistics, but destroys object structure and higher-order correlations) to end up with a so called spectrally matched noise version, see Fig. 5.

This strategy was used to investigate how neurons in the secondary visual cortex of macaques respond to naturalistic textures versus spectrally matched noise (J. Freeman et al., 2013; Ziemba et al., 2024). The study found that while primary visual cortex neurons responded similarly to both types of stimuli, V2 neurons exhibited stronger responses to naturalistic textures. This suggests that V2 is more attuned to higher order statistical structures in natural scenes, beyond just spectral content.

Spectrally matched noise was also used to explore how neurons in the cat primary visual cortex adapt their receptive fields when exposed to natural scenes versus

(Sharpee et al., 2006). It resulted that neural filters adaptively change with the input ensemble to increase the information carried by the neural response about the filtered stimulus. This adaptation enhances sensitivity to under-represented frequencies, aligning with optimal encoding theories.

Alternatively, one can match higher-order statistics using models like the texture synthesis algorithm from Portilla and Simoncelli, which does a very visual cortex like description of the different visual features and higher order correlations (Portilla & Simoncelli, 2000). In principle, it captures information at multiple scales and orientations via a complex wavelet transform, analyses correlations between coefficients at adjacent positions, orientations, and scales and then adjusts a random noise image in several iterations until it matches the measured statistics. A notable finding is that V1 neurons generally respond to low-level statistics but are not aware of higher-order arrangements that V2 neurons detect. In a systematic study, they recorded from macaque V1 and V2 neurons using synthetic texture images that were generated by matching local texture statistics up to a certain order to natural images (Ziemba et al., 2016). They found that despite the fact that those textures shared luminance, contrast and orientations, V2 neurons could reliably distinguish them, whereas V1 neurons treated those textures more similarly. In fact, V2 showed selectivity for naturalistic texture patterns, and tolerance to changes that preserved those statistics, implying that V2 extracts types of texture features that V1, being better at encoding local features, is oblivious about. Consistent with this, in another study it was found that macaque V4 has a preference for specific higher order statistics combinations (Okazawa et al., 2015). The mouse higher cortical area LM has also shown a higher selectivity than V1 for higher order textures (Bolaños et al., 2024).

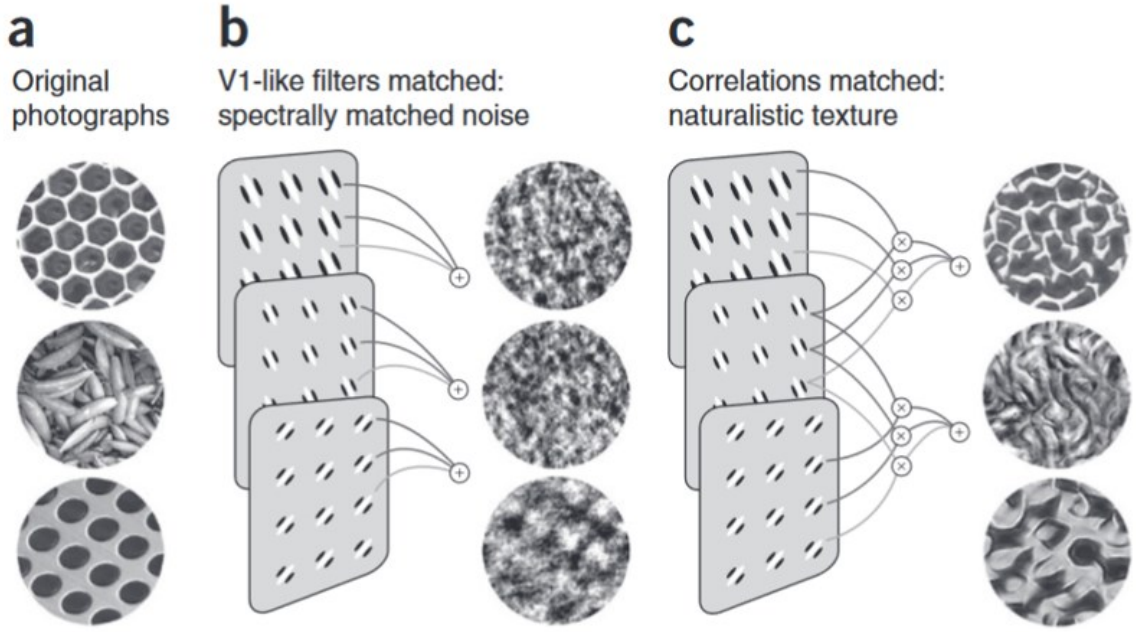


Figure 5. Example pipeline for generating spectrally matched noise stimuli.

(**a**) Original images. (**b**) The images go through filters and energy filters that have different orientation and spatial frequency tuning (similar to a population of V1 neurons. The outputs are noise images that retain the same average orientation and frequency structure across space but lose higher order features. (**c**) Naturalistic texture images with more naturalistic features than the original image is synthesized. [Figure reproduced from Freeman et al., 2013, Nature Neuroscience, © Springer Nature. Used with permission for academic purposes.]

In humans, this approach has shown useful in explaining the trade of between the ability to distinguish between different stimuli and the ability to recognize the same stimulus under varying conditions in peripheral vision (Ziemba & Simoncelli, 2021). The study suggests that the visual system dynamically balances these two abilities, meaning that responses to structured images like natural scenes may favor higher selectivity in central vision while prioritizing invariance in the periphery and since spectrally matched noise removes high-order statistical structures, the visual system may struggle to form a stable representation of it, treating it as a more ambiguous or unstructured input.

Alternatively, people have looked for higher order patterns in adjacent image pixels. Several studies have been using versions of these stimuli across different species and have pointed out that there is an increased perceptual sensitivity to these higher order features (Hermundstad et al., 2014; Tkačik et al., 2010; Victor et al., 2017; Yu et al., 2015). However, there is too large of a parameters space to be tested and only limited aspects of natural scenes can be probed by these local pixel patterns

representations. So, covering this large combinatorics while having no possibility to relate these responses to global distributions of orientations or spatial frequencies is of limited benefit.

Overall, while these alternative stimuli have revealed a lot about the role of higher visual areas in texture perception and global percept, lower level feature processing and the role of V1 therein still remain understudied (Tesileanu et al., 2022). Another point that these studies are touching but, as things stand, cannot close in on is how local features converge towards a global perception. Again, this would first require a better understanding of the role of V1 in this processing path. Furthermore, the hierarchical interactions between different visual areas also need to be captured well. One would need experimental designs that address aspects such as modulation of the responses of a neuron based on the context found in the area that surrounds the receptive field of this neuron, shown for simple and naturalistic stimuli (Adesnik et al., 2012; Haider et al., 2010; A. J. Keller, Dipoppa, et al., 2020; A. J. Keller, Roth, et al., 2020; Kirchberger et al., 2021; Pecka et al., 2014; Vinje & Gallant, 2000; Williford & Von Der Heydt, 2016).

1.3.2 Studying natural vision with complex and composite textures.

Iterative optimizations are not the only way of matching nature statistics to synthetic stimuli for the purpose of studying processing of higher order statistics. Even more complex stimuli can be generated by using supervised learning (Cadena et al., 2019). Textures coming from the lower layers of deep convolutional neural networks (dCNN) preserve local edge related statistics, whereas higher layers synthetize the more global structures. Presenting such stimuli to monkeys revealed that these multi-step nonlinear transformations can do a very good job in predicting V1 computations, which would highlights how some of V1's nonlinear computations are analogous to processes from a deeper layer of a dCNN (Miao & Tong, 2024), so, this, along with the superficial layer like edge detection, make it analogous to a multi-layer artificial neural network.

The implementation of such methods has brought with it a new way of approaching the questions at hand. The dCNNs are trained on the responses of neurons to nature scenes with the aim of optimizing the prediction of what drives the neuronal responses

(Walker et al., 2019). Then, stimuli that match this optimization are generated. These stimuli come from a rich, non-bias, data driven probing and they do indeed contain within them aspects of first and higher order statistics of nature, but their complexity makes it challenging to interpret the specific features that drive the neuronal responses. Also, these optimized stimuli might not necessarily represent naturally occurring visual signals, which means they need to be paired with a thorough ecological validation. For example, characterizing neurons in terms of basic feature (orientation, spatial frequency) tuning, then reverse correlating their responses to the preferred natural images to find and compare their favorite visual patches and the respective MEIs that emerge in terms of the first and higher order statistics.

Further types of complex stimuli that mimic natural scene structure or composite images have been used to test models of vision. Metamers are physically different images that produce the same responses in a model of early vision, to get insight on what information is lost at each stage (J. Freeman & Simoncelli, 2011). Human observers could not distinguish the original natural images from their respective metamers in the periphery, which suggests that early visual pathway is selective of the statistics of the input they retain and discard.

These studies emphasize that visual processing of natural input cannot be understood by studying isolated stimuli alone, but that textures and composites are crucial for uncovering the reaction of cortical neurons to the joint statistics of features in natural scenes. Though useful within the aim of predicting neuronal responses to visual stimuli, these methods come with risks and limitations (Bowers et al., 2022). Deep convolutional neural networks can mimic certain hierarchical aspects of human visual processing, but they lack hierarchical correspondence with the brain visual areas and cannot capture particularities related to recurrent feedback mechanisms, top-down modulation, context and predictions. They are also not made for dynamic processes and temporal integration. Furthermore, and very important in this specific context, their vulnerability to adversarial perturbations or unexpected variations emphasizes their lack of ability to generalize statistical irregularities the way a brain can. So, in aspects where the model receives too sparse training, it could miss out on meaningful patterns or even build misleading blind spots.

1.3.3 Studying natural vision with visual noise.

Looking at all the non-linearities and operations that neurons have employed to complex stimuli, it becomes apparent that natural scenes engage many mechanisms at once. A way to get around this convoluted signal, is by using simplified noise stimuli in order to look at the contributions of different types of nonlinearities.

The classic visual noise experiment is correlating random pixel flickers with neuronal spikes. This is called spike-triggered average (STA) and is used to estimate the neuron's receptive field. However, neurons are not purely linear but often have multiple selectivity subunits and potentially even a mixture of linear and nonlinear subunits (Ding et al., 2023). A good workaround for this was looking at the spike-triggered covariance (STC) instead. For example, studies have applied this analysis to responses of macaque V1 neurons with two main stimulus dimensions in their complex cells: one dimension with even-symmetric Gabor correspondence and one with odd-symmetric Gabor correspondence (Chen et al., 2007; Rust et al., 2005). Also, this helped getting insight into the orthogonally suppressive dimensions that connect to the existence of a cross orientation suppression that is seen in the monkey (Vogels & Orban, 1990), cat (DeAngelis et al., 1992; Priebe, 2016) and mouse (Katzner et al., 2011; Niell & Stryker, 2008) studies. And ultimately, the findings from using this approach helped explain the phase of position invariance previously seen in V1 neurons, since multiple subunits with similar orientation tuning but shifted position could sum over position or phase to generate a robust response. (Chen et al., 2007).

A more complex noise stimulus used to probe nonlinear integration is structured noise. This stimulus was used to compare responses to naturalistic noise versus binary checkerboard noise to see if neurons do a simple summation of patch luminance (Touryan et al., 2005). This was not the case as although natural image patches with the same local contrast were shown, neurons responded more sparsely and selectively to the naturalistic noise compared to random inputs. This relates to how natural images are able to access the complex subfields and inhibitory surrounds which are not activated by white noise randomness. Seeing how neurons integrate information nonlinearly over a larger area, noise stimulus with certain second-order

correlations such as a background orientation and motion would be more successful in revealing contextual tuning mechanisms.

Natural visual scenes exhibit a predominance of low spatial frequency content and a broad distribution of orientations (Field, 1987; Girshick et al., 2011). Using higher order texture stimuli would help to fill the gap in understanding the neuronal representation of visual feature diversity in natural scenes. Randomized broadband noise patterns have been used for bandwidth-processing adjacent questions in monkey (Goris et al., 2015), cat (Ladret et al., 2023) or mouse (Han et al., 2022) studies. Overall, these studies suggest the presence of diverse encoding modes, whereby highly selective neurons do a temporally stable encoding of single features (for example stimuli with a sharp distribution in orientation space), whereas less selective neurons encode the mixture of features (for example stimuli with a broad distribution in orientation space). Since several types of tuned units are engaged, slow recurrent interactions gradually fine-tune the representation of the most salient feature. This provides a robust framework for natural context discrimination and perception. Certainly, the mixture of stimulus features also increases the population recruitment due to the increased presence of features like orientations or spatial frequencies which this pool of neurons is tuned to (Goris et al., 2015; Ladret et al., 2023; Niell, 2015; Niell & Stryker, 2008; Roth et al., 2012).

In practical terms, visual noise stimuli have helped build more accurate models of V1 and extend previous understanding of complex cells with additional excitatory or inhibitory subunits and overall V1 computations. For example, modeling work has shown that a model that includes multiple linear filters feeding into a nonlinear unit can do a better job than a linear-nonlinear (LN) model in predicting V1 responses (Vintch et al., 2015). This helps us see V1 neurons through a feature combination lens which is crucial for understanding encoding of natural scenes, which are rich in combinations. However, much has to be resolved about the low-level features that are most effective in driving cortical network activity and the contribution of their intrinsic tuning versus feedback or recurrent elements, surround suppression etc. Also, we lack insight into the perceptual and behavioral extent of this rich information representation.

1.3.3.1. Random phase textures with modified feature bandwidth.

Seeing the open gap in a deeper understanding of how the rich feature variability of nature is represented in the visual cortex we sought to find a fitting stimulus model with visual noise stimuli of higher order textures and tailorable feature distribution. Motion Cloud is a stimulus package that generates a parametrically tunable, dynamic subset of random phase textures. The central motion component is characterized by spectral envelopes which concentrate on a plane (Leon et al., 2012), (see Fig. 6).

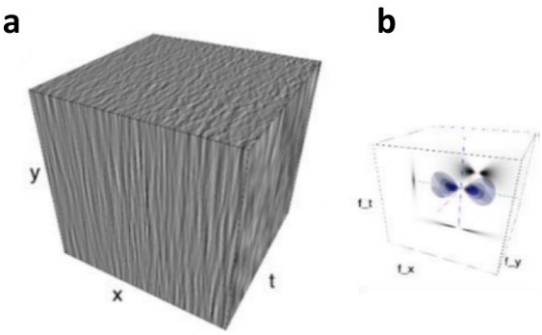


Figure 6. Example Motion Cloud movie rendering.

Shown as 3-dimensional display of all the composing frames and respective Fourier spectrum. (**a**) Motion Cloud movie with an isometric view of the 3D faces of a. Dimensions x (height) and y (width) are shown on the first frame (x, y, time (t) = 0). Top of the cube (x= 0, y, t) shows the progression of x over movie frames while right facet (x, y, t = 0) shows y progression over the movie frames (**b**) The corresponding Fourier energy spectrum (amplitude normalized to 1) in its isometric view to show the projections onto three orthogonal planes. The package is open to access on GitHub: https://github.com/NeuralEnsemble/MotionClouds

This means, the movies that are generated here have continuous and coherent motion. Furthermore, each frame contains information about luminance at position (x, y), and the distribution of luminance over time, as shown in Figure 1. This model takes advantage of the the first order motion conservation of phase shuffling related information in the Fourier domain (Derrington et al., 2004). This allows for different cube facets in Figure 1a, while keeping an almost identical Fourier energy spectrum in Figure 1b. Thus, drifting stimuli of differing spatial frequencies can be generated (Nixon & Aguado, 2012). Furthermore, the generative model uses different variables to express the full set temporal scale (fixed) versus individual features (each feature has image specific relative amplitude). This allows for a reliable generation of multiple parametrized sets of movies, i. e. stimuli with different orientation bandwidths and sets with different spatial frequency bandwidths. Other texture stimuli exist that mimic natural motion by matching the statistical structure of natural scenes and higher-order motion features, but they are more suitable for studying mid-level processing (Portilla & Simoncelli, 2000). Motion Cloud stimuli are parametrically controlled and allow

precise manipulation of individual visual features, making them ideal for probing early visual processing.

1.4 What is still missing in our understanding of natural vision?

Historically, visual neuroscience has had its main focus on individual neurons and their operations. But, ultimately, perception of natural scenes is generated by coordinated populations of neurons. Thus, the ideal description of natural signal representation would consist of a global view of population and areal interactions, including a description of the feedforward and feedback contributions at specific input-output hierarchy levels of the signal path. Furthermore, the contributions of this processing in the behavioral state and output of the animal and vice versa, to see how the task context (e.g. a more difficult task might require employment of different mechanisms than a simple one) influences processing.

1.4.1 Population level processing and remaining gaps.

Employing high yield methods such as 2-photon calcium imaging, Neuropixels electrode recordings etc., have facilitated observation of large-scale population responses. Thanks to this, we now know that V1 population responses to natural stimuli are highly dimensional (Stringer et al., 2019). In fact, the dimensionality of the responses scaled roughly with the number of neurons recorded, meaning larger populations contribute to more independent feature combinations, as each neuron adds some unique information. From a decoding perspective, this means that V1 population output can be used to reconstruct or recognize natural images, and since machine learning can do that, probably so can the downstream visual areas (Froudarakis et al., 2014). Next important point in the difference in the difference between population code and single-neuron coding, is the degree of expression. While a single neuron can at most modulate their response preferentially for different stimuli (stay silent for some, fire more for others, etc.) a population can employ combinatorial patterns to create unique codes for each stimulus, which is more robust, less sensitive to errors from individual units and better suited for distinguishing overlapping representations. Moreover, population recordings show that neurons are not simply

independent encoders, but recipients of a collective modulation that is corroborated by the presence of noise correlation or ensemble activity fluctuations. Overall, how attention or how the previously mentioned global population dynamics affect natural scene processing remains unstudied. There are hints that anesthesia has a hindering effect on the sparsening of the responses and noise correlations and surround suppression (Bharioke et al., 2022; A. J. Keller, Roth, et al., 2020; Lee et al., 2021), which suggests that brain state and behavioural context gates the efficiency of the population code.

1.4.2 The concrete input-output hierarchy: subcortex, layers and areas.

We've looked for the computations in the biology now we turn our attention to the biology of the computations. V1 has been the spotlight of this introduction, mainly due to its crucial role in basic feature processing that is relevant to look into for inferring the amount of natural statistics represented in the brain. However, vision involves a cascade of visual areas, starting from the conversion of light photons to electrical signals in the retina, to subcortical areas, layer specifically to V1 and then further along a feature relevant streams of HVAs (e.g. V2, V4, IT stream for object recognition in the monkey, though for the mouse this division is not as clear cut). Each downstream level extracts more complex statistics. But what intermediate representations are built, at which level, what contributes to their establishment and how much the original statistics do they retain and pass through? The path from edge processing to object perception is very complex but also very informative because it involves broader interactions between multiple cortical areas (and inevitably, layers), some it also being feedback, modulatory signals.

Furthermore, at each stage there is a battery of actors (neurons) that play different roles. For example, in the V1 neuron surround suppression scenario we have strong indications that key players are SOM interneurons which inhibit (similarly tuned) excitatory neurons upon a large uniform stimulus context and VIP interneurons which serve as a potential disinhibition switch upon contrasting contexts (Adesnik et al., 2012; A. J. Keller, Dipoppa, et al., 2020; Kirchberger et al., 2021). Whether this is a specific mechanism for this type of operation or a universal mode for input predictability encoding, is not yet known.

So what kind of experiment could help us build a unified view of natural processing in the visual system? First, we look at an intermediate feature that could help reveal complex computations along the visual pathway: a stimulus not too simple but still simple enough that one can infer its attributes, like for example the richness in feature information, a nature statistics attribute that might find its representation in the visual system. Investigate its influence on the population activity on V1, L2/3 of awake mice (via high spatial resolution 2-photon imaging) to look at the contribution of behavioral state and attention on the possible effect. Also run this experiment with anaesthetized mice to see how much global or cortico-cortical connections influence this effect. Follow with simultaneous recordings from multiple areas to see the global activity (for example with Widefield imaging) and potential downstream areas involved, then do causal perturbations of these downstream areas to see if there is also feedback modulation on the V1 activity. If that is the case, look at the contributions of key neuron players (pyramidal, SOM and PV) simultaneously (for example by delivering cell specific opsins of different emission color, while imaging over several color channels) to build a circuitry. Also look at their responses in the focal downstream area. Finally, break down the contributions of feedforward and feedback inputs by looking at the encoding in L4 (along with immediately related subcortical areas, e.g. with Neuropixels electrode recordings) versus L2/3 (plus somato-dentritic activity correlation) and deeper layers (especially L5 if anesthesia has an interesting effect and especially L6 if behavioral state has an effect). In this study, we've laid a good foundation towards that comprehensive experimental suite.

1.4.3 A link to perception and behavior.

Despite having well characterized neural codes, the relation between these and the perception and behavioral output of the animal is often missing. When does a certain pattern of V1 responses correspond to seeing a particular thing? To build this bridge between coding and perception, neural recordings need to be coupled with animal behaviour tracking along with engaging naturalistic stimulus detection or discrimination tasks. CNN-based models have been shown to significantly outperform traditional LN models and Gabor filter bank models in predicting V1 neuron responses to natural images (Cadena et al., 2019), which suggests that V1 might itself be optimized for more than just local statistical encoding and might be influenced by the

tasks the animal needs to perform. Thus, another important addition to the experimental suite presented in the previous section, would be combining the effect relevant experiments with mice which are being trained on different levels of recognition or discrimination difficulty and compare how task engaged responses differ from passive viewer responses, how these responses change over expertise levels and how the difficulty of the task at hand might call for different coding strategies.

As we fill the remaining gaps, we move closer to figuring out how the visual system makes sense of the natural world with such ease, and how different ecological contexts or environmental challenges are defining the choice of encoding strategies.

1.5 How does this study enrich the field?

The visual system has evolved to optimally represent the visual attributes of nature. One dynamic aspect of nature, that despite its high potential for revealing important coding mechanisms has received little attention henceforth, is the variability of feature distributions. We consider the encoding of this variability not only a puzzle to solve, but also see it as a tool for parsing roles and contributions of different levels of the visual system in the collective representation and perception of nature regularities.

To address the puzzle aspect of this variability, we wanted to generate random phase texture movies of varying orientation and spatial frequency bandwidths: very narrow (quasi grating-like) and mid distributions representing occasional scene specific cases and the broad bandwidth stimuli representing the overall distributions of these features in nature typically. Thus, it can be said that our stimuli would become more naturalistic as their bandwidths would broaden. We wanted to investigate different aspects of natural feature distribution representation in the mouse visual system through a wide selection of methods (see methods section for general principles behind each method).

1.5.1 Specific aims.

I. What effects does feature bandwidth enrichment have on the neuronal responses of mouse V1?
   a. Is it feature specific, i.e. is there a difference between orientation and spatial frequency?
   b. Is the effect simply due to intrinsic tuning properties of V1 neurons?

To address this, we aimed to perform 2 photon imaging for a population-wide, high spatial resolution perspective into the responses of the overall neuronal population (injecting a calcium indicator that is expressed in most cells) as well as specifically pyramidal neurons (using a genetically modified mouse model that only expresses the calcium indicator in excitatory neurons) of mouse V1 L2/3. Orientation or spatial frequency bandwidth tailored stimuli were to be presented in a full-field screen which covered a broad visual field of our awake passively viewing mice*. Based on literature and fundamental response tuning principles of V1 neurons, an increase in population recruitment with the feature bandwidth enrichment was expected.*

II. Does behavioral state/arousal have an effect on the feature enrichment effect?

We also aimed to track the behavioural state of the animals while showing the passive stimuli to have an insight whether the neuronal responses are influenced by the behavioral or arousal state of the animal. *Given the passive viewing context of our visual experiment, we expected a behavioral state independent stimulus enrichment effect on the neuronal responses.*

III. Is there a focal point in the visual cortex that drives the bandwidth enrichment responses?

With the help of widefield imaging, we wanted to look at the global cortical activity across several visual cortex areas in search of a focal HVA zone involved in the representation of broadband stimuli. *Given the unpredictability*

and higher order textures embodied in broadband orientation stimuli, we expected a lateralized focal point for their representation.

IV. Is there a difference in feedforward and feedback activity in this context?

For this, we planned to use Neuropixel probes to carry out electrophysiological recordings at deeper cortical layers and reaching as far as the optical part of the superior colliculus (SC). *We expected strong responses, especially for a very information rich mixed stimulus that contains broadband orientations and broadband spatial frequencies.*

V. Does stimulus bandwidth enrichment have a perceptual effect in the mouse?

Finally, we wanted to investigate the perceptual correlate of orientation and spatial frequency bandwidth enrichment by running a visual discrimination task in a mouse touchscreen chamber. To avoid confounding effects, we planned to let the mice discriminate upon central spatial frequencies when enriching in orientation bandwidth and to let them discriminate upon central orientation when enriching in spatial frequency bandwidth. *Based on the neuronal responses we expected to see improved discrimination for orientation bandwidth enrichment.*

2. RESULTS

2.1 Broadening the orientation bandwidth of a visual stimulus to match nature statistics leads to an increased recruitment and response amplitude in mouse V1 neurons.

In regards to its visual aspect, nature offers a wide diversity of feature combinations, with edges following a conditional dependence that very much relates to the structural context of a spatial area. Depending on the targeted area, an observer could face a very uniform scene in terms edge orientations (Fig. 7a, region 2), or something of a richer scene that spreads more in terms of the distribution of edge orientations present (Fig. 7a, region 1). But overall, there is a very broad distribution of orientations (orientation bandwidth) present in natural scenes (Fig. 7a, nature library) and much is left to be discovered about the visual cortex neuronal responses to these context dependent orientation bandwidths.

Therefore, we set out to explore mouse V1 responses to different orientation bandwidth stimuli. We designed several sets of motion cloud stimuli with the intention of approximating the range of orientation bandwidths in natural images (Ganguli & Simoncelli, 2014; Girshick et al., 2011; Tkačik et al., 2011) (Fig. 7b). We started with a narrow-range stimuli with a (5°) bandwidth that is meant to simulate uniform scenes like region 2, then broadened the orientation bandwidth of the motion cloud stimuli to a mid-range level of (25°), and finally the most broadband stimuli had a (45°) orientation bandwidth to best match the dominant orientation bandwidths of natural images (Fig. 7a) (Girshick et al., 2011).

To avoid texture specific effects, several motion cloud stimuli were generated for each bandwidth. Additionally, all stimuli were kept at a set central orientation of 0° and a set central spatial frequency of 0.04 cpd to match the response tuning of neurons in mouse visual cortex (Andermann et al., 2011; Marshel et al., 2011; Roth et al., 2012). Each individual stimulus consisted of a 5-s long, rightward drifting movie (at 1Hz temporal frequency).

---

* málà is a native Szechuan province seasoning consisting of Szechuan pepper corn and chili. Literally translated it means numbing (麻- má) and spicy (辣 – là).

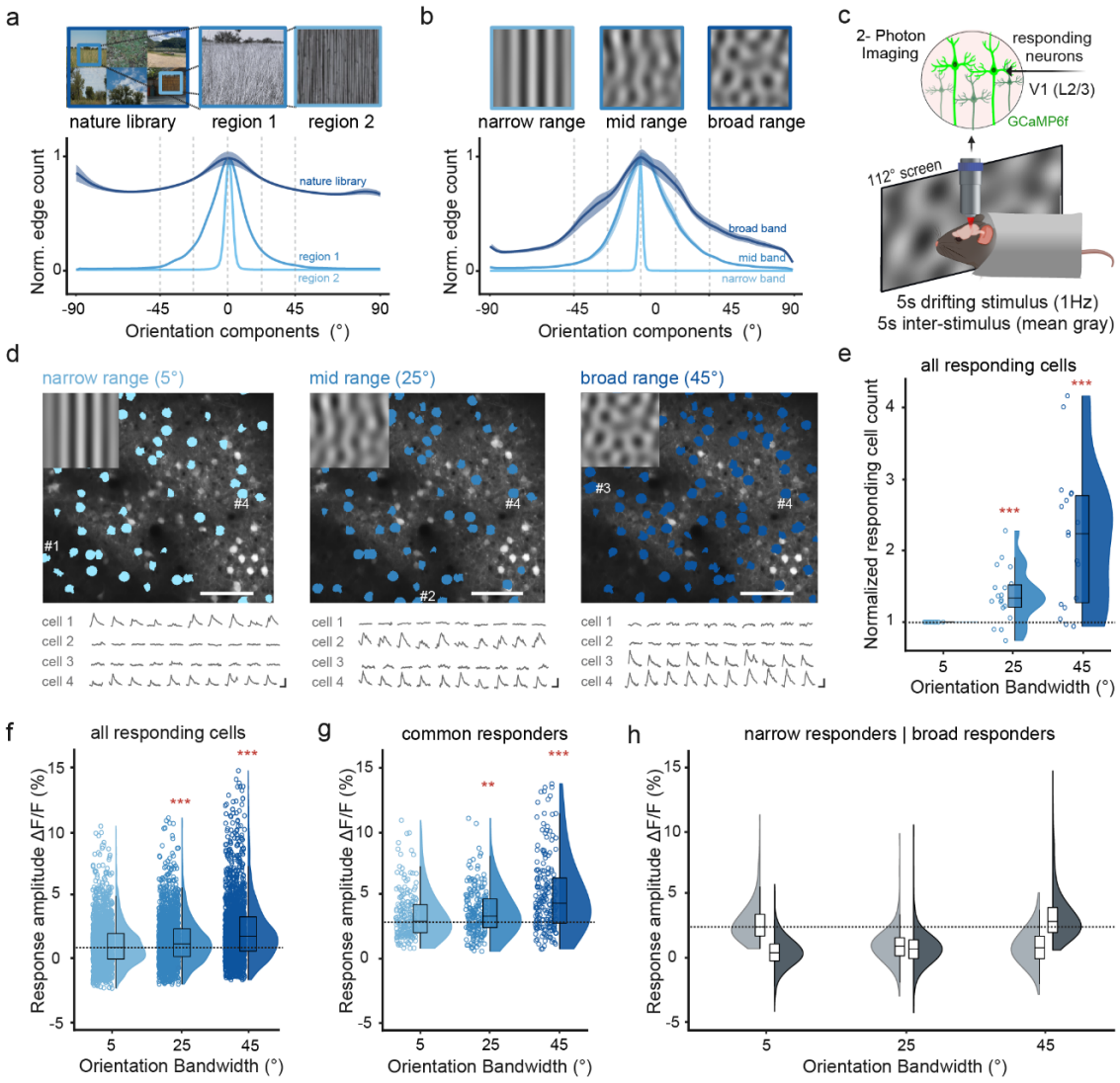


Figure 7. Broad orientation bandwidth stimuli drive strong and ample responses in the mouse V1.

(**a**) Mean orientation distributions from natural images (example images adapted from Tkacik G et al, "Natural images from the birthplace of the human eye ", PLoS ONE 6: e20409 (2011), n = 1077 images, shading shows s.e.m.). Single-cell selectivity was tested at specific orientations (dashed lines, Fig. 11, Fig. 12). Regions 1, 2 show mid-range or narrow orientation bandwidth examples. (**b**) Same as a) for broadband orientation bandwidth stimuli (5°, 25°, and 45°, n = 300 frames each). (**c**) Awake, head-fixed mice viewed differing orientation bandwidth stimuli while 2-photon imaging V1 neural activity in layer 2/3. (**d**) Example session, showing V1 neural responses to different orientation bandwidths (5°, 25°, 45°, blue masks show responsive cells, scale bar: 100 µm, corresponding stimuli at top left). Traces show single-trial responses of four example cells with different orientation bandwidth preferences (scale bar: 1 s, 20% ΔF/F). (**e**) Counts of responsive neurons to each orientation bandwidth, normalized by the number of cells responding to the 5° bandwidth (n = 18 sessions from 9 mice). (**f**) Difference between mean response amplitude versus baseline for each orientation bandwidth. Shown are neurons with a significant response for at least one condition (1394 out of 4892 neurons, 29% ± 2% neurons per session; mean ± s.e.m.). Horizontal line shows the median narrow response amplitude. (**g**) Same as f) for neurons that consistently responded to all bandwidths (12% of responding cells per session, 252 total neurons). Horizontal line shows the median narrow response amplitude. (**h**) Same as f) for neurons that preferentially responded to narrow-band (left violin plots, 7% of responding cells per session, 152 in total), or to broadband (right violin plots, 23% of responding cells per session, 487 neurons in total). Box plots indicate median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Dotted line is the narrow-band median. Stars mark significant (Bonferroni correction for two tests, α = 0.025) differences from two-sided tests against narrow condition: LME test (panel e) Wilcoxon signed rank (panels f, g) Wilcoxon ranked sum (panel h). Panel c was created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

The whole battery of different orientation bandwidth stimuli was presented in a pseudo-random order (with 5-s mean-gray inter-stimulus interval) to awake, head-fixed mice. The mice were passively sitting in a tube, to reduce motion related artefacts and activity as much as possible. The specific cortical target at this step was the mouse V1, layer 2/3, as a lot of visual input signal integration and interesting response nonlinearities happening there, which could be expected of an information enriching stimulus such as broadband orientation stimuli.

With high spatial resolution population overview being a strong criterion, we opted for doing 2-photon imaging of V1, layer 2/3 to measure the neural responses of neurons in V1, ubiquitously expressing the calcium indicator GCaMP6f (Fig. 7c, d). To evaluate the impact of different orientation bandwidths on the number of responding neurons, we first used a one-sided Mann-Whitney U test ($p < 0.05$) to identify all neurons with a significant response for each stimulus condition against baseline. Where possible, we used a linear mixed effects (LME) model, which accounts for nested data from multiple animals, to test for significant differences across conditions. Then, we followed with a Bonferroni correction for the number of performed tests (see Methods).

Broadening the orientation bandwidth, first more than doubled the number of visually-responsive neurons across all sessions (Fig. 7e), which means recruitment of V1 mouse neurons goes up as orientation bandwidth in increased (normalized responding cell count, both tested against the narrow stimulus: mid range (25°) = 1.38 ± 0.09, LME model test against narrow, T = 4.64, $p = 4.9 \times 10^{-5}$; broad range (45°) = 2.2 ± 0.23, LME model test against narrow, T = 5.58, $p = 3.04 \times 10^{-6}$; mean ± s.e.m, n = 18 sessions from 9 mice. See also per mouse comparison in Fig. 8a.

Percentage of visually responsive neurons over number of total neurons per session was: narrow range = 11% ± 1%; mid range = 15% ± 2%; broad range = 21% ± 2%. Beside an increase in total number of responding neurons, expanding the orientation bandwidth also strongly increased neuronal response amplitudes, which peaked for the broadest orientation bandwidth (Fig. 7f, $\Delta F/F_{narrow}$ = 1.16% ± 0.04%; $\Delta F/F_{mid}$ = 1.44% ± 0.04%, Wilcoxon signed rank test against narrow, $p = 5.21 \times 10^{-12}$; $\Delta F/F_{broad}$ = 2.29% ± 0.06%, Wilcoxon signed rank test against narrow, $p = 5.25 \times 10^{-66}$, mean ± s.e.m., n = 1394 neurons from 18 sessions, 9 mice, see also per mouse comparison in Fig. 8b).

Thus, expanding orientation bandwidth, increased both the neuronal recruitment as well as and amplitude of stimulus-responsive V1 neurons. Furthermore, this increase in response amplitude was also seen in neurons that consistently responded to all orientation bandwidths (Fig. 7g), called common responders. These neurons significantly responded to all orientation bandwidth stimuli (one-sided Mann-Whitney U test, $p < 0.05$) but strengthened their response amplitudes to the broadest orientation stimuli (Fig. 7g, $\Delta F/F_{narrow}$ = 3.39% ± 0.14%; $\Delta F/F_{mid}$ = 3.67% ± 0.13%, Wilcoxon signed rank test against narrow, $p = 2.6 \times 10^{-3}$; $\Delta F/F_{broad}$ = 5.05% ± 0.22%, Wilcoxon signed rank test against narrow, $p = 2.1 \times 10^{-9}$, mean ± s.e.m., n = 252 neurons from 18 sessions, 9 mice). This indicates that the widespread signal response gain was not merely due to a larger neuronal recruitment but specifically by a general increase in neural responsiveness.

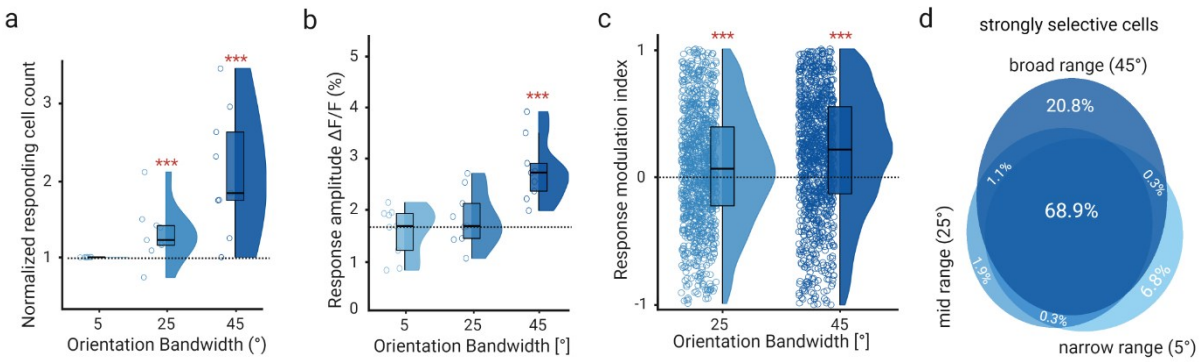


Figure 8. The orientation bandwidth effect is consistent across animals.

(**a**) Counts of responsive neurons to each orientation bandwidth, normalized by the number of cells responding to narrow (n = 9 mice). (**b**) Responses amplitudes for all neurons with a significant response for at least one stimulus (2096 responsive neurons out of 4892 neurons in total). (**c**) Response modulation index (RMI) for neurons with significant positive responses to any enriched stimulus (RMI = (enriched – narrow)/ (enriched + narrow)). Mid band modulated neurons: 57% positively (RMI = 0.39 ± 0.01, mean ± s.e.m.), 43% negatively (RMI = -0.33 ± 0.01, mean ± s.e.m). Broadband modulated neurons: 67% positively (0.20 ± 0.01, mean ± s.e.m), 33% negatively (-0.18 ± 0.01, mean ± s.e.m.). (**d**) Distribution of neurons that selectively respond to a single or multiple orientation bandwidth stimuli or common responders that significantly responded to all stimulus bandwidths. The percentage of neurons from all neurons that significantly and selectively respond to either of the bandwidth stimuli is shown (366 neurons from 18 sessions in 9 mice). For visualization only, outliers were excluded from distributions in panels a-c. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). In panels a and b stars show significance for the Wilcoxon signed rank test against narrow responses, dotted line is narrow median. In panel c stars show significance for an LME model test against zero, dotted line is zero. The significance threshold was adjusted upon Bonferroni correction for number of tests performed (α = 0.025, 2 tests, panels a-c). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

To assess for presence of broadband-selective neurons and see if they contribute to this increased neural responsiveness we calculated a bandwidth selectivity index (which is the response difference between a preferred orientation bandwidth and the mean of the response to the other two bandwidths, divided by their sum) and used a shuffle control (see Methods) to identify bandwidth-selective neurons. Most of the resulting bandwidth-selective neurons preferred the broadband orientation stimuli, and a small subpopulation of neurons preferred narrow band stimuli (Fig. 8d). By this logic, additionally recruited neurons would be picked due to their orientation tuning preference being away from the orientations available in the narrow band stimulus. Whereas neurons tuned to the narrow band orientation content would become unresponsive. However, since only few neurons were significantly bandwidth-selective (n = 25, 7, 76 neurons for the narrow, mid, and broad range, respectively, versus 252 common responders) in comparison, this increased responsiveness to broadband stimuli is expected to be driven by more than just broadband-selective neurons.

In the end, we compared the very selective neuronal populations to see if selectivity could have an impact on response enrichment. Specifically, we looked at the response amplitudes of narrow selective neurons versus broadband selective neurons (Fig. 7g) using a less conservative method for assigning the bandwidth-selective label to all neurons that only significantly responded to either the narrow- or to broadband condition (one-sided Mann-Whitney U test, $p < 0.05$). Being highly selective, both populations had their strongest responses to their respective bandwidth, which they are selective for (Fig. 7h, maximum response mean ± s.e.m.: $\Delta F/F_{narrow}$ 3.36% ± 0.28%, n = 152 neurons, $\Delta F/F_{broad}$ = 3.60% ± 0.17%, n = 487 neurons; Wilcoxon ranked sum test for narrow versus broad, $p = 0.089$).

Enriching the orientation bandwidth of a stimulus drives a widespread V1 layer 2/3 activity by recruiting a large number of neurons that respond very strongly.

2.2 Broadening the spatial frequency of the stimulus does not have a strong impact in the responses of mouse V1 layer 2/3.

Natural scenes are rich in spatial frequencies, with coarse structures dominating over fine structures. Thus, spatial frequency distributions in nature have substantially more power in the low-frequency bands compared to the high-frequency bands (Ganguli & Simoncelli, 2014; Wiesbrock et al., 2022) (Fig. 9a). V1 neurons of mice have a limited range of spatial frequency tuning which have already been characterized for (Girshick et al., 2011; Marshel et al., 2011; Roth et al., 2012).

We wanted to see how broadening the spatial frequency bandwidth of random phase texture, to approximate what could be found in different natural scenes (Fig. 9a, b) affects neuronal responses in V1. For this, we picked a limited range of bandwidths to broaden upon on top of the most preferred V1 population spatial frequency (0.004 cpd, 0.04 cpd, and 0.4 cpd bandwidths; on top of a central spatial frequency of 0.04 cpd). We presented the stimuli in a pseudo-random order while measuring the activity of V1 neurons as described in the previous section. (Fig. 9c, d).

In contrast to orientation bandwidth, broadening the spatial frequency bandwidth did yield an increase in the number of visually responsive neurons (Fig. 9d, e; normalized responding cell count both tested against narrow condition: mid range (0.04 cpd) = 1.03 ± 0.1, LME model test against narrow: T = 0.35, p = 0.72; broad range (0.4 cpd) = 1.3 ± 0.24, LME model test against narrow, T = 1.43, p = 0.16; n = 16 recordings from 9 mice. See also per mouse comparison in Fig. 10a. Percentage of responsive neurons over total neurons per session: narrow range = 7% ±1%, mid range = 7% ± 2%; broad range = 9% ±1%, mean ± s.e.m.).

Furthermore, we could not see a significant increase in the neural response amplitudes to the broad versus narrow spatial frequency bandwidth stimuli for all responsive neurons (Fig. 9f; $\Delta F/F_{narrow}$ = 1.7% ± 0.1%; $\Delta F/F_{mid}$ = 1.9% ± 0.1%; Wilcoxon signed rank test against narrow: p = 0.09; $\Delta F/F_{broad}$ = 1.9% ± 0.1%; Wilcoxon signed rank test against narrow: p = 0.25, mean ± s.e.m, n = 888 neurons from 16 sessions, 9 mice, see also per mouse comparison in Fig. 10a).

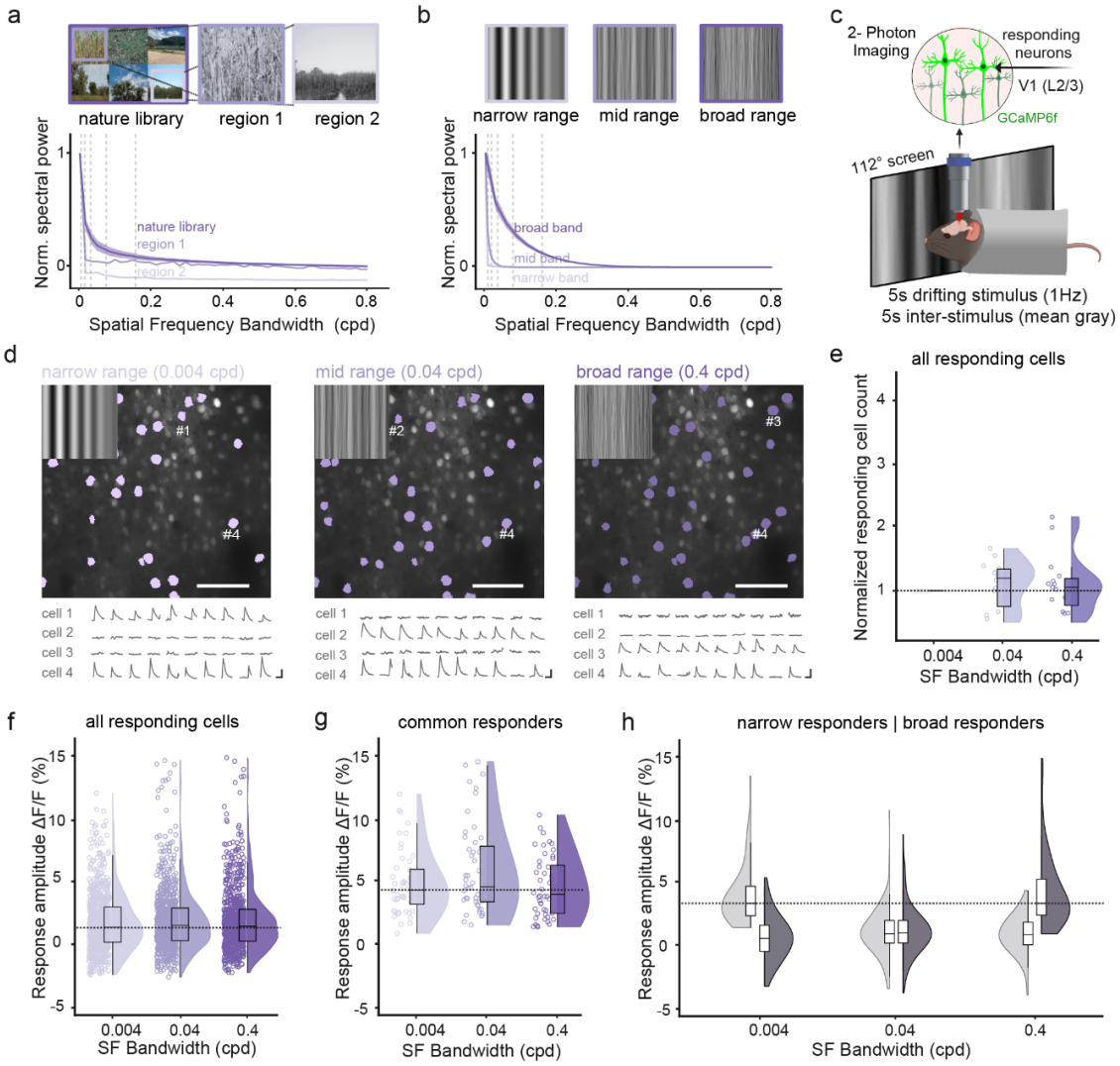


Figure 9. V1 neurons maintain the same overall response despite spatial frequency bandwidth richness.

(**a**) Mean spatial frequency distributions from natural images (example images adapted from Tkacik G et al, "Natural images from the birthplace of the human eye ", PLoS ONE 6: e20409 (2011), n = 1077 images, shading shows s.e.m.). Single-cell selectivity was tested at specific spatial frequencies (dashed lines, Fig. 11). (**b**) Same as a) for broadband stimuli with different spatial frequency distributions (0.004 cpd, 0.04 cpd, and 0.4 cpd, n = 300 frames each). (**c**) Awake, head-fixed mice viewed differing orientation bandwidth stimuli while 2-photon imaging V1 neural activity in layer 2/3. (**d**) Example session, showing V1 neural responses to different spatial frequency bandwidths (0.004 cpd, 0.04 cpd, 0.4 cpd, purple masks show responsive cells, scale bar: 100 µm, corresponding stimuli at top left). Traces show single-trial responses of 4 example cells with different spatial frequency bandwidth preferences (scale bar: 1 s, 20% ΔF/F). (**e**) Counts of responsive neurons to each spatial frequency bandwidth, normalized by the number of cells responding to 0.004 cpd bandwidth (n = 16 sessions). (**f**) Difference between mean response amplitude versus baseline for each orientation bandwidth. Shown are neurons with a significant response for at least one condition (603 of 3185 neurons, 19% ± 2% neurons per session; mean ± s.e.m.). Horizontal line shows the median narrow response amplitude. (**g**) Same as f) for neurons that consistently responded to all bandwidths (68 neurons, 9% of responding cells per session). Horizontal line shows the median narrow response amplitude. (**h**) Same as f) for neurons that preferentially responded to the narrow-band stimulus (left violin plots, 24% of responding cells per session, 133 in total), or the broadband stimulus (right violin plots, 26% of responding cells per session, 132 neurons in total). Box plots indicate median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Dotted line is the narrow-band median. No significant (Bonferroni correction for two tests, α = 0.025) differences from two-sided tests against the narrow condition: LME test (panel e) Wilcoxon signed rank (panels f, g) Wilcoxon ranked sum (panel h). Panel c was created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

Similarly, for the common responders there were also no differences in response amplitude across different spatial frequency bands (Fig. 9g, $\Delta F/F_{narrow}$ = 4.9% ± 0.4%; $\Delta F/F_{mid}$ = 6% ± 0.5%, Wilcoxon signed rank test against narrow: p = 0.11; $\Delta F/F_{broad}$ = 4.6% ± 0.4%, Wilcoxon signed rank test against narrow: p = 0.71, mean ± s.e.m., n = 68 neurons).

The simplest explanation for this difference in response enrichment between the two features could be the neuronal tuning (Roth et al., 2012). If V1 neurons are broadly tuned to many spatial frquencies, they will already be responding at full capacity to the narrow spatial frequency stimulus and therefore the recruitment is not happening as we broaden the spatial frequency bandwidth. To address this, we looked at the bandwidth selectivity differences between both features to see if more unselective neurons are found across all spatial frequency bandwidth stimuli. Turns out that actually the opposite of that happens, with a larger population of common responder neurons being found in response to the different orientation bandwidths (Fig. 7g & 9g). Furthermore, significantly bandwidth-selective neurons (selected against a shuffle control) were more common for broadband spatial frequency stimuli than for broadband orientation stimuli (11.1% versus 7.7%, respectively), which goes against the premise of an unselective population with globally broad spatial frequency tuning.

Another interesting point was that the mid-range stimuli which are between the narrow and broad spatial frequency stimuli and contain elements from both, turned out to have the smallest subpopulation of selective neurons. An equally large population of neurons was selective for the other two ends of the spectrum (Fig. 10d). This suggests that the cortical network maintains a homeostasis of activity in response to broadening of spatial frequency bandwidth, by balanced up and down regulation of responses while maintaining a stable number of recruited neurons and general response amplitudes.

When comparing the activity of the highly selective neurons for narrowband or for broadband spatial frequency stimuli, we saw similar response amplitudes (Fig. 9h, maximum response mean ± s.e.m $\Delta F/F_{narrow}$ = 4.32% ± 0.33%, n = 133 neurons, $\Delta F/F_{broad}$ = 4.99% ± 0.45%, n = 132 neurons; Wilcoxon ranked sum test for narrow versus broad: p = 0.52). This is in line with the stable representation idea.

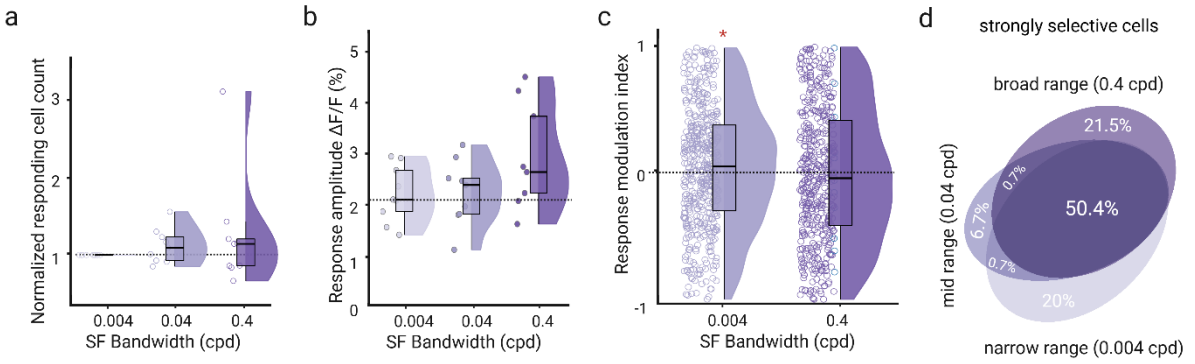


Figure 10. Spatial frequency bandwidth enrichment doesn't impact responses across mice.

(**a**) Counts of responsive neurons to each spatial frequency bandwidth, normalized by the number of cells responding to narrow (n = 9 mice). (**b**) Responses amplitudes for all neurons with a significant response for at least one stimulus (n = 9 mice, 888 responsive out of 3185 total neurons). (**c**) Response modulation index (RMI) for neurons with significant positive responses to any enriched stimulus (RMI = (enriched – narrow)/ (enriched + narrow)). Mid band modulated neurons: 54% positively (0.40 ± 0.02, mean ± s.e.m), 46% negatively (-0.37 ± 0.02, mean ± s.e.m). Broadband modulated neurons: 48% positively (0.30 ± 0.02, mean ± s.e.m), 52% negatively (-0.18 ± 0.02, mean ± s.e.m). (**d**) Distribution of neurons that selectively respond to a single or multiple spatial frequency bandwidth stimuli or common responders that significantly responded to all stimulus bandwidths (135 neurons from 16 sessions in 9 mice). For visualization only, outliers were excluded from distributions in panels a-c. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). In panels a and b stars show significance for the Wilcoxon signed rank test against narrow responses, dotted line is narrow median. In panel c stars show significance for an LME model test against zero, dotted line is zero. The significance threshold was adjusted upon Bonferroni correction for number of tests performed (α = 0.025, 2 tests, panels a-c). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

To check whether single neurons had a different global response width to orientations versus the spatial frequencies within the ranges we used in the motion cloud stimuli, we did extra recordings where we showed simple gratings with different orientation and spatial frequency combinations (Fig. 11). We tested responses to the orientation and spatial frequency ranges that were used in our stimuli. There ranges were inspired nature (dashed lines in Fig. 7a, 9a). The neurons showed varied tuning to the different gratings, but their global tuning widths matched the range of orientations and the range of spatial frequencies used. This suggests that the lack of stronger responses to broadband spatial frequency stimuli was not due to overly broad tuning stopping specific subpopulations from being recruited. Instead, the stronger response to broader orientation bandwidths seems to be a unique feature of V1 neurons and does not apply to spatial frequency.

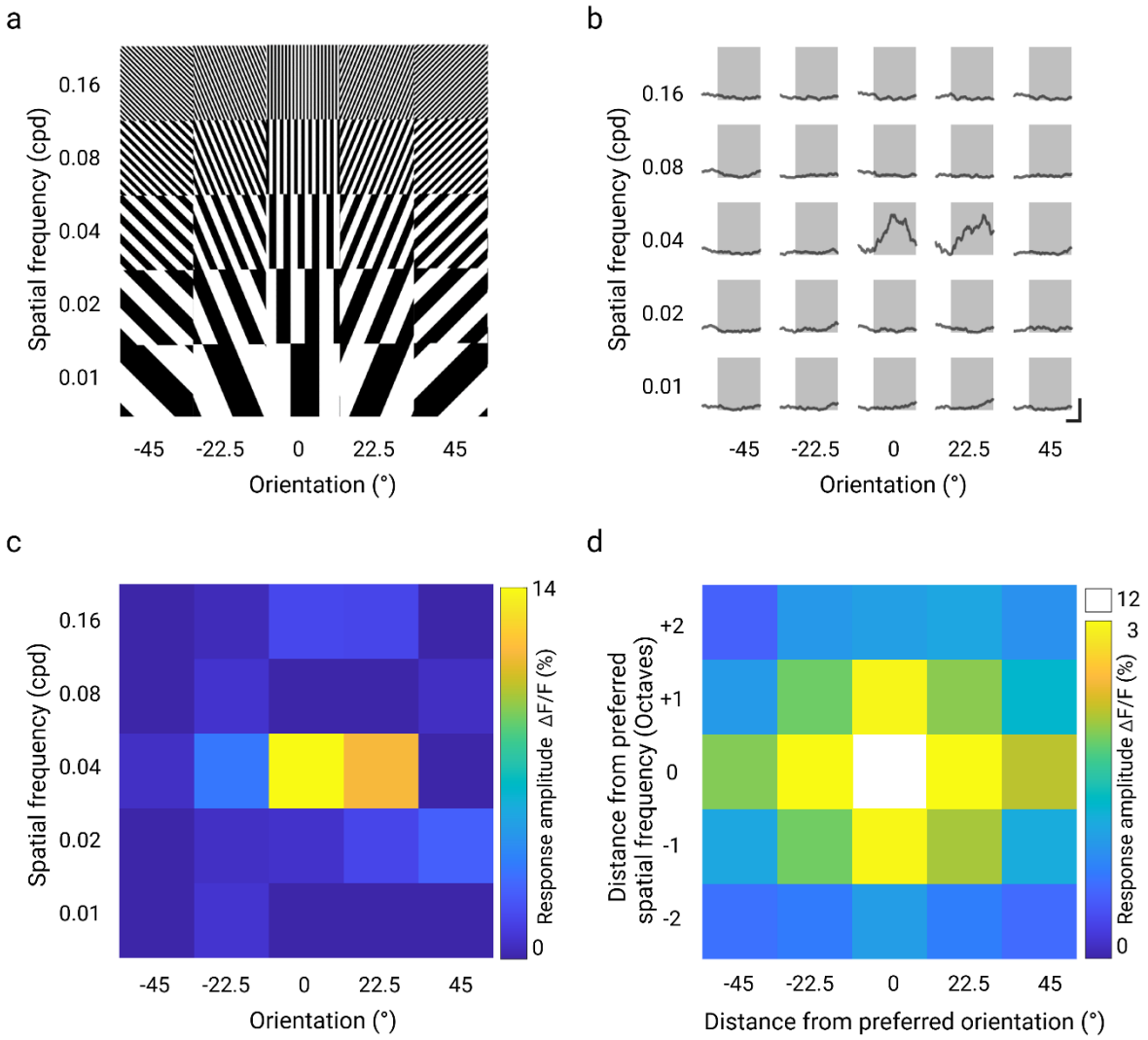


Figure 11. V1 tuning is comparable across our tested range or orientations and spatial frequencies.

(**a**) Overview of 25 combinations of visual grating stimuli, consisting of 5 different orientations (-45°; -22.5°; 0°; 22.5°; 45°) and spatial frequencies (0.01 cpd; 0.02 cpd; 0.04 cpd; 0.08 cpd; 0.16 cpd). All stimuli were presented as pseudorandom sequences to 3 awake mice in 6 sessions in total. (**b**) Mean response traces of an example V1 neuron for each stimulus combination. The shaded area is the stimulus duration. (Scale bar: 1s, 20% ΔF/F). (**c**) The average response strength to each stimulus combination for the neuron is shown in panel b. (n = 20 presentations per stimulus). Visible is a narrow orientation and spatial frequency tuning. (**d**) Averaged tuning specificity for all responsive V1 neurons (n = 971 cells). The response matrix was shifted for each neuron so that the peak orientation and spatial frequency responses were in the center. Datapoints that were not measured because they were outside of our tested range were set to NaNs and not weighted. X- and Y-axes, therefore, show the distance from the preferred orientation and spatial frequency to measure the width of the population orientation and spatial frequency tuning within the presented stimulus range. Paired Wilcoxon signed rank test of the full-width half maximum (FWHM) of a Gaussian fit for the central X-axis (for orientation) and Y-axis (for spatial frequency) showed a small but significant difference between the two tuning widths: $FWHM_{orientation} = 1.76 \pm 0.04$ bins, $FWHM_{spatial\ frequency} = 1.67 \pm 0.04$ bins (Wilcoxon signed rank test, $p = 1.91 \times 10^{-38}$). This suggests that V1 neurons were slightly less selective for different grating orientations. Importantly, this is the opposite of what would have been expected if the selective increase in neural responses to broad orientation bandwidth had been solely caused by additional recruitment of V1 subpopulations with very narrow orientation tuning. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

2.3 A Gabor filter model that includes surround suppression can accurately predict responses to broadband orientations.

A simple explanation for the stronger neural responses to broader orientation bandwidths could be that more neurons are activated when their preferred orientation falls outside the narrow stimulus range. The presentation of new orientations in the broader bandwidth stimuli then simply recruits these neurons. However, our previous findings challenge the expectation that this is the sole contributing element. For example, we saw increased response amplitudes in neurons that responded to both narrow and broad stimuli when the orientation bandwidth was expanded (Fig. 7h), and we didn't observe the same effect for broadband spatial frequency stimuli.

To explore this further, we built a detailed Gabor filter model based on known receptive field properties of V1 neurons (Fig. 12a, Fig. S1 and see Methods). The model used Gabor filters spanning the same range of orientations and spatial frequencies as our visual stimuli (Fig. S3). We also collected new data by recording responses to both narrow and broad orientation bandwidth stimuli, while also measuring the preferred orientation of each neuron. Then, we compared the response amplitudes to narrow versus broad orientation stimuli, grouped by each neuron's preferred orientation (Fig. 12b, c).

As expected, neurons tuned to the central orientation of 0° responded most strongly to the narrow bandwidth stimuli, showing a clear link between orientation tuning and response. In contrast, responses to the broadband orientation stimuli did not show this tuning specific pattern (Fig. 12b, c). Our Gabor filter model captured this general trend in neural responses based on tuning (yellow lines show best model fits).

However, the simpler "tuning only" model overestimated how much responses depended on orientation, especially for 0° tuned neurons responding to narrow bandwidth stimuli (Fig. 12b, c; $R^2_{\text{tuning only}} = -1.266$). We thought this mismatch might be due to the more regular spatial structure of narrow stimuli, which could lead to stronger surround suppression in neural responses.

To test this, we expanded the model by adding surround suppression: responses from the image center were influenced by surrounding Gabors with the same orientation (Fig. S2). This updated "suppression" model reflected the real orientation dependence

of neural responses to both narrow and broad stimuli more accurately, and it provided a much better fit to the data (red lines show best-model fits; $R^2_{suppresion} = 0.584$).

We also calculated an “orientation response modulation index” (ORI RMI) to compare how neurons responded to narrow versus broad orientation bandwidth stimuli. This index was defined as the difference between the responses to narrow and broad stimuli, divided by their sum. As expected, neurons with their preferred orientation further away from the central 0° orientation showed the largest increase in ORI RMI. (Fig. 12d, ORI $RMI_{-45°} = 0.12 \pm 0.04$, T = 2.79, $p = 0.01$; ORI $RMI_{-22.5°} = 0.08 \pm 0.04$, T = 2.11, $p = 0.04$; ORI $RMI_{0°} = -0.03 \pm 0.04$, T = -0.90, $p = 0.37$; ORI $RMI_{22.5°} = 0.10 \pm 0.04$, T = 2.4, $p = 0.02$; ORI $RMI_{45°} = 0.13 \pm 0.04$, T = 3.74, $p = 10^{-3}$; mean ± s.e.m., LME model test against zero, n = 11 recordings from 5 mice).

Matching the better fit to response amplitudes (Fig. 12b, c), only the surround model accurately predicted how neural response amplitudes changed with broadband orientation stimuli across all orientation tunings (Fig. 12d, red line, $R^2_{suppresion} = 0.395$). In contrast, the simpler tuning-only model predicted a stronger drop in response for neurons tuned to 0°, which was not supported by the actual data (Fig. 12d, yellow line $R^2_{tuning\ only} = -1.67$). Similar to what we saw for response modulation, we also found that more neurons were recruited when their orientation tuning was further from the central 0° (Fig. 12e). To include this effect in the model, we linearly scaled the model fit to response modulation, assuming that stronger responses would lead to more neurons being activated. Once again, the surround suppression model matched the data more closely than the tuning-only model ($R^2_{tuning\ only} = 0.563$, $R^2_{suppresion} = 0.889$).

As seen in our earlier recordings (Fig. 9), broad spatial frequency bandwidth did not affect neuronal response amplitudes, regardless of the neurons’ orientation tuning (Fig. 13a, spatial frequency response modulation index (SF RMI): $SF\ RMI_{-45°} = -0.06 \pm 0.06$, T = -1.37, $p = 0.19$; $SF\ RMI_{-22.5°} = -0.06 \pm 0.05$, T = -1.38, $p = 0.18$; $SF\ RMI_{0°} = -0.12 \pm 0.07$, T = -1,37, $p = 0.08$; $SF\ RMI_{22.5°} = -0.03 \pm 0.09$, T = -0.41, $p = 0.68$; $SF\ RMI_{45°} = -0.01 \pm -0.06$, T = -0.17, $p = 0.87$, mean ± s.e.m., LME model test against zero, n = 11 recordings from 5 mice). Similarly, we found no change in the number of visually responsive neurons upon spatial frequency broadening and no clear link to the orientation tuning of these neurons (Fig. 13b).

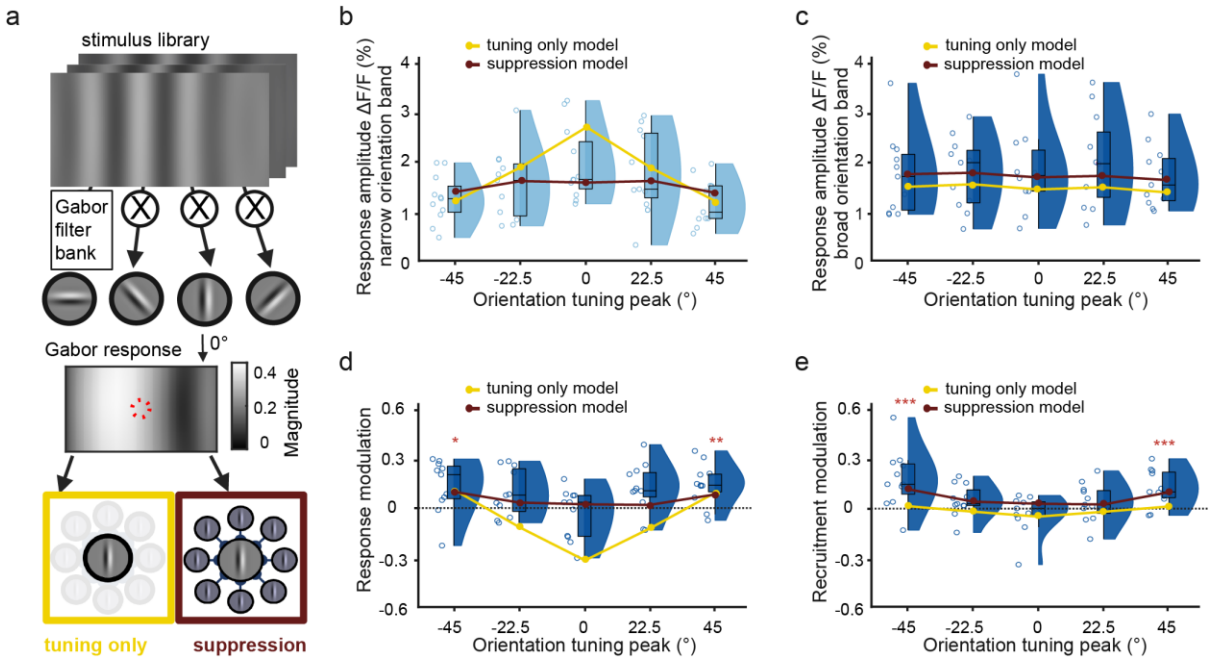


Figure 12. Gabor filter model links orientation tuning and surround suppression to broadband responses.

(**a**) Schematic of a Gabor filter bank to model neural responses. Gabor filters of different orientations and spatial frequencies were used to jointly predict the measured responses to the presented broadband stimuli (see Fig. S1 and S2 for details). Model predictions were either based on the responses of a central Gabor alone ("tuning only" model; bottom left, yellow box) or by additionally accounting for the impact of surround suppression ("suppression" model; bottom right, red box) (**b**) Mean response amplitude towards narrow bandwidth stimuli, depending on the orientation tuning of V1 neurons (n = 11 sessions in 4 mice). Neurons in each session were binned based on their orientation tuning to a set of full-contrast sinewave gratings. Yellow circles and line show the best-fitted response of the tuning-only model. Red circles and line show the best fit of the suppression model. (**c**) Same as b) for broad orientation bandwidth stimuli. Both the tuning only and the suppression model were fit to narrow and broad orientation bandwidth stimuli at the same time. (**d**) Mean response modulation of orientation-tuned cells by expanding the orientation bandwidth (derived from responses in b and c). Yellow circles and line show the predicted response modulation from the tuning-only model. Red circles and line show predicted response modulation from the suppression model. The horizontal black dotted line at 0 indicates no bandwidth modulation. (**e**) Same as panel d) but for recruitment modulation of orientation-tuned cells by expanding the orientation bandwidth. The ORI recruitment modulation index was calculated as the difference between the numbers of responsive neurons to narrow and broad orientation bandwidth stimuli divided by their sum. Yellow and red circles indicate fits of the tuning only and surround suppression model, respectively. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Stars in panels d and e mark significant differences against 0 based on an LME model (two-sided test). The significance threshold for each test was α = 0.01 after Bonferroni correction for five individual tests. Panel a was created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

Taken together, these results suggest that both orientation tuning and surround inhibition play key roles in shaping broadband response modulation in V1. Notably, the model predicted that narrow bandwidth stimuli, with their more regular spatial structure, should trigger stronger surround suppression. This could explain how the broadband orientation stimuli lead to increased neural responses due to a reduction in surround suppression.

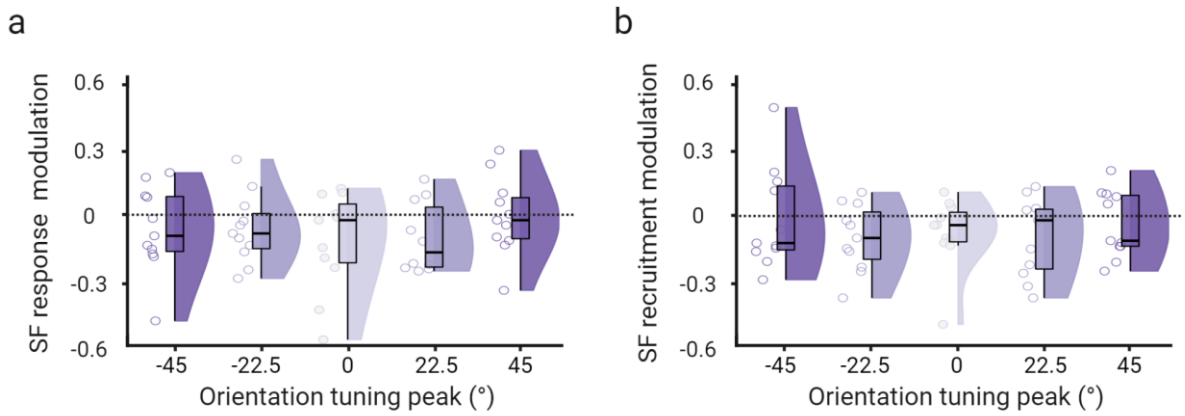


Figure 13. Broadband spatial frequency response modulation is not linked to the orientation tuning.

(**a**) Mean response modulation of orientation-tuned cells by expanding the spatial frequency bandwidth (n = 11 sessions in 4 mice). Neurons in each session were binned based on their orientation tuning within the range of the broad orientation bandwidth stimulus. The horizontal black dotted line at 0 indicates no bandwidth modulation. As in our earlier recordings (Fig. 9), expanding the spatial frequency bandwidth did not have an impact on neuronal response amplitudes, irrespective of their orientation tuning (Spatial frequency response modulation index (SF RMI): SF RMI-45° = -0.06 ± 0.06, T = -1.37, p = 0.19; SF RMI-22.5° = -0.06 ± 0.05, T = -1.38, p = 0.18; SF RMI0° = -0.12 ± 0.07, T = -1,83, p = 0.08; SF RMI22.5° = -0.03 ± 0.09, T = -0.41, p = 0.68; SF RMI45° = -0.01 ± -0.06, T = -0.17, p = 0.87, mean ± s.e.m., LME model test against zero, n = 11 recordings from 5 mice). (**b**) Same as panel a) but for recruitment modulation of orientation-tuned cells by expanding the spatial frequency bandwidth. The SF recruitment modulation index was calculated as the difference between the numbers of responsive neurons to narrow and broad SF bandwidth stimuli divided by their sum. We also found no changes in the recruitment of visually-responsive neurons or a clear dependence on their orientation tuning in panel b. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Stars in all panels mark significant differences against 0 based on an LME model. The significance threshold was α = 0.01 after Bonferroni correction for performing five tests. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

2.4 Surround modulation plays a key role in accurately predicting how individual neurons respond to broadband stimuli.

To directly verify how orientation tuning and surround modulation influence broadband responses in individual neurons, we conducted additional experiments using awake 2NiellJ mice expressing GCaMP6s in excitatory neurons (Fig. 14a). We presented a series of narrow and broadband motion cloud stimuli, all centered at a spatial frequency of 0.04 cpd and a 0° orientation. The “narrow” condition used motion clouds with a tight spatial frequency bandwidth of 0.004 cpd and a narrow orientation bandwidth of 5° orientationn bandwidth (0.004 cpd and 5°, respectively).

The “SF” condition used stimuli with a broad spatial frequency bandwidth (0.4 cpd) and a narrow orientation bandwidth (5°). The “ORI” condition had the opposite: a narrow spatial frequency (0.04 cpd) and a broad orientation bandwidth (45°). The “mixed” condition combined both broad spatial frequency and broad orientation bandwidth (0.4 cpd and 45°) to test their combined effect on neural responses (Fig. 14b). We presented these stimuli either as full-field, covering the entire screen, or limited to the center of the receptive field area for most recorded neurons in each session. The center stimulus was 15° in size, chosen to match the optimal receptive field size of layer 2/3 V1 neurons (A. J. Keller, Roth, et al., 2020) (see also Fig. 15 and Methods). The average measured RF size was 30.45° ± 2.04°, which confirmed that the center stimulus was consistently placed within the receptive field of the selected neurons.

To explore how higher-order features differed across stimulus conditions, we measured salient orientations and stimulus regularity or linear predictability. Specifically, the question was: how predictable the is the center based on the surround information? We quantified this predictability by using raw coefficient correlation (Portilla & Simoncelli, 2000) (Fig. 14c). The narrow stimulus shows the highest level of spatial auto-correlation and hence predictability. Broadening the spatial frequency bandwidth still maintained the elongated structure of the underlying gratings. In contrast, broadening the orientation bandwidth produced irregular, lattice-like patterns that were less redundant than the clear, elongated edges.

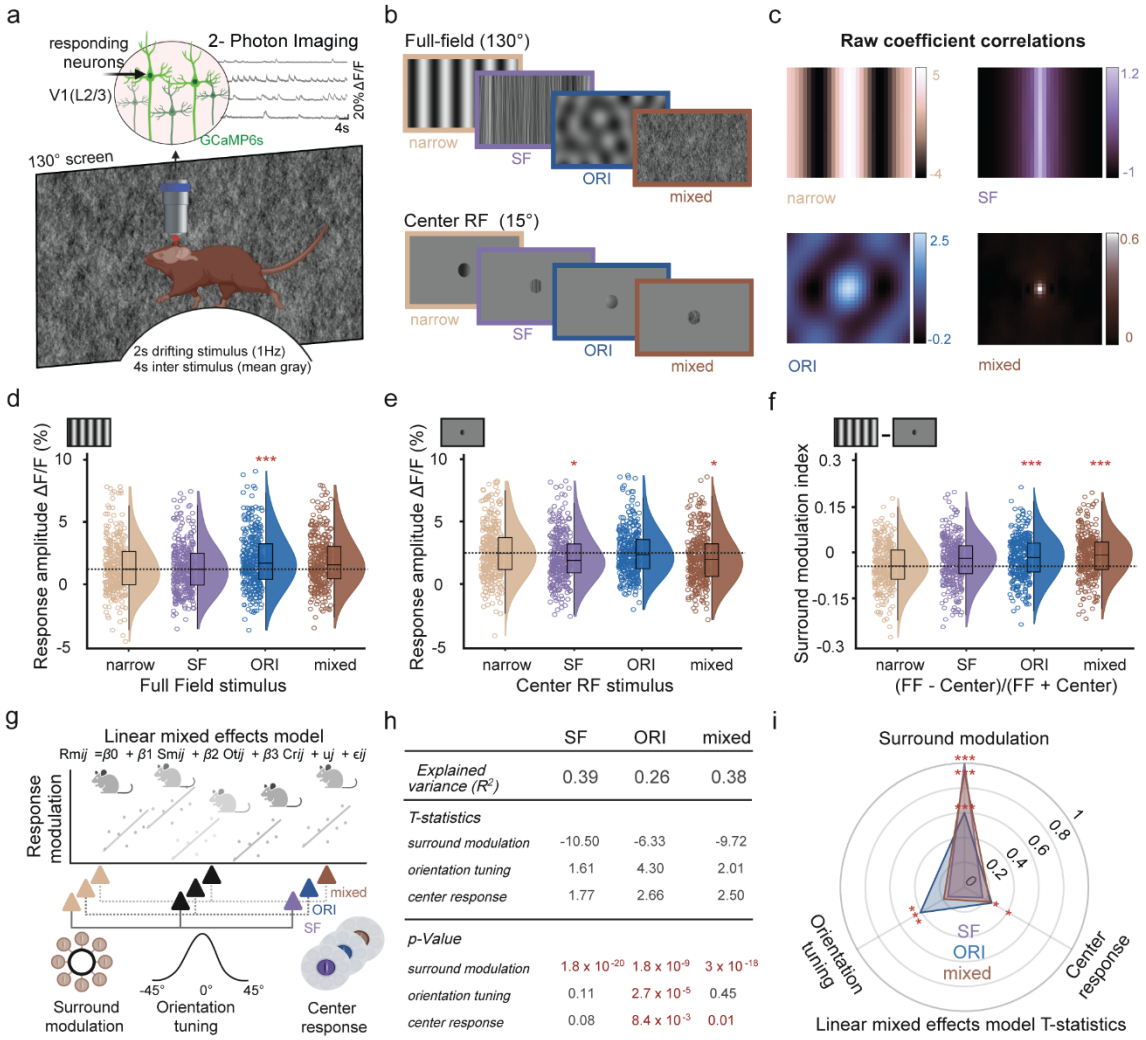


Figure 14. Broadband visual stimuli modulate neural responses by reducing surround suppression.

(**a**) Visual stimulation and imaging setup. (**b**) Example frames from each stimulus category. Stimuli had a central spatial frequency of 0.04 cpd and orientation of 0° and shown full-field or through a 15° circular aperture, centered on the receptive field of imaged neurons. Frame colors indicate the condition identity throughout the figure. (**c**) Raw coefficient correlations were computed as the central samples of the spatial auto-correlation of lowpass filtered stimuli. (**d**) Mean response amplitude difference versus baseline for each full-field stimulus condition. Shown are all neurons with a receptive field at the center stimulus and significant responses to any of the center stimulus conditions (378 out of 8210 neurons). The horizontal line shows the median for the narrow stimulus. (**e**) Same as panel d) but with responses to each center stimulus condition. (**f**) Surround modulation index for individual neurons, calculated as the difference between full-field responses in d) minus center stimulus response in e) divided by their sum. The horizontal line shows the median for the narrow stimulus condition. (**g**) Illustration of the linear mixed effects (LME) model with multiple regressors. Surround modulation, orientation tuning, and center responses were used to predict single-cell broadband modulation for each condition. The model also considers the mouse identity for each neuron (see Methods). (**h**) Explained variance of the full LME model for each stimulus condition and T statistics and *p* values for each regressor. Significant regressors are marked in red. (**i**) Polar plot to compare T statistics results for each stimulus condition. Each color-coded triangle shows the normalized T values for each regressor, indicating its respective contribution to the model prediction. Across all conditions, surround modulation was the most important predictor. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Stars mark significant (Bonferroni correction for three tests, α = 0.0167) differences from two-sided tests against the narrow condition: Wilcoxon signed rank (panels d-f), LME test (panel i). For visualization only, outliers were excluded from distributions. Panels a and g were created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

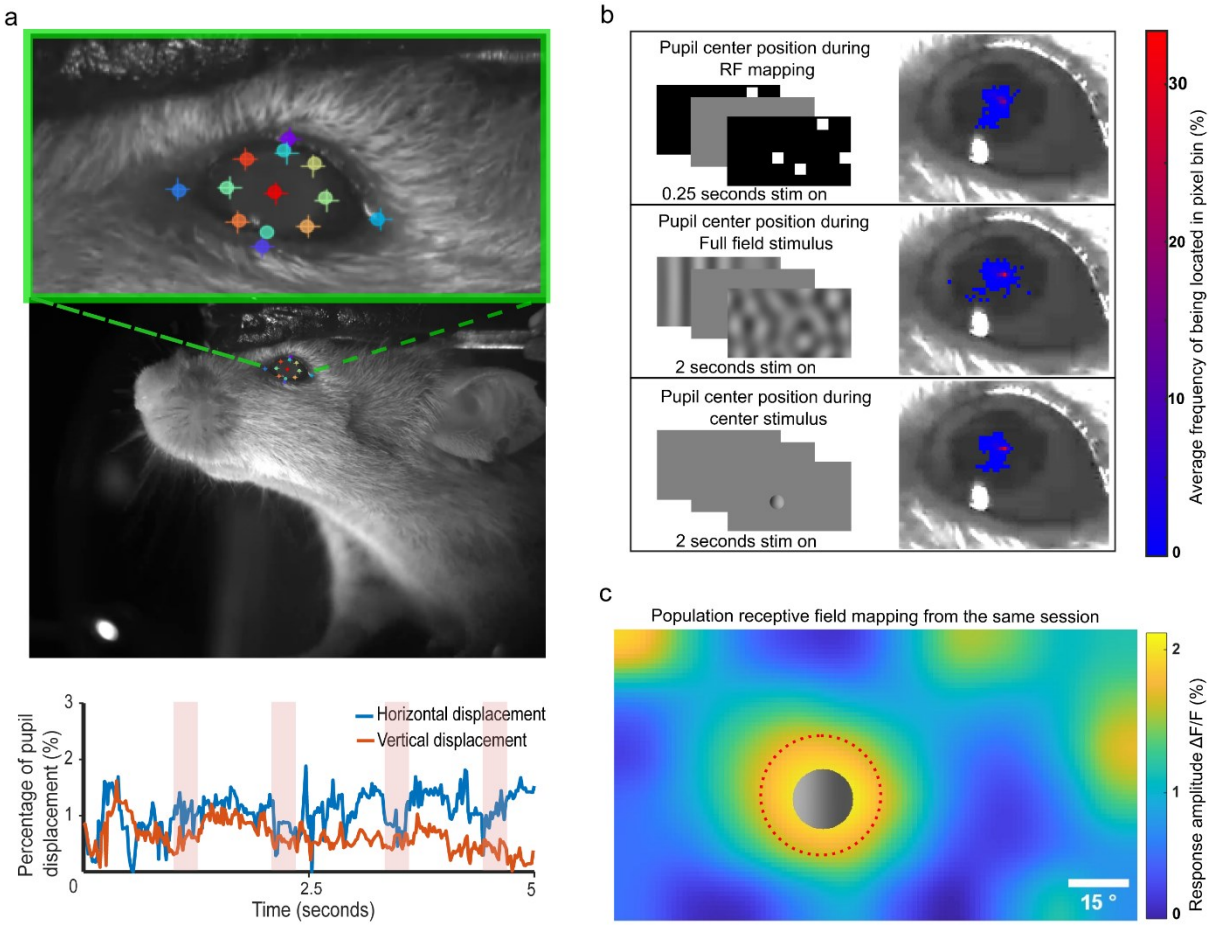


Figure 15. Pupil movements do not influence neural responses to stimuli presented at the RF center.

(**a**) Example frame from the behavior camera. Expanded on top is the mouse eye, labeled at several locations to train a DeepLabCut[1] model that followed the eye position throughout the session. We labeled the top, bottom, left, and right eyelid edges to obtain a reference of the outer range to where the pupil could move. Then, we tracked the pupil location (eight reference points around the pupil edge and one at the center). The bottom panel shows example pupil movements during a randomly selected 5-second time window during the receptive field (RF) mapping. The pupil displacement is shown in pixels, normalized by the total possible displacement range for each axis (horizontal or vertical). The pink shaded areas mark the sparse noise stimulus presentations that were used for RF mapping (0.25 seconds stimulus duration with 1-second inter-stimulus intervals on a mean gray screen). (**b**) Mean heatmap of the center pupil locations during either the RF mapping (top), full-field stimulus presentations (middle), and the center stimulus presentations (bottom). For visualization purposes, we used 2x2 pixel binning of pupil positions relative to the camera resolution to calculate the fraction of frames containing the pupil center in each pixel bin ("location frequency"). This was done separately for each stimulus presentation condition (all RF stimuli, all full-field stimuli, or all center stimuli). For each condition, we found that the pupil remained largely invariant at the eye center with a very restricted surrounding area with low position frequency. This shows that changes in pupil position are not a major confound for interpreting neural responses during the different stimulation paradigms. (**c**) RF mapping of the center responding cells from the same example session. The RF map was matched to the screen size, and an example center stimulus is shown at the center of the RF center on the screen. The scale bar is 15°. Average RF size (dotted red circle surrounding the stimulus) of the population across all the sessions was 30.45° ± 2.04° (mean ± SD), ensuring that the center stimulus was placed within the receptive field of the "center responding" cells, shown in Fig. 14. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

When both spatial frequency and orientation bandwidth were broadened in the mixed condition, most regular structures disappeared, causing a strong drop in how predictable the center was based on the surrounding stimulus. Moreover, in the mixed

condition we also found the highest reduction in higher-order image structure (Fig. S3, computed as coefficient magnitude statistics (Portilla & Simoncelli, 2000)). Higher-order structures were also increasingly present in the ORI condition, which might help explain its saliency to the V1 layer 2/3 population.

As in our earlier findings, neural responses were stronger for broadband orientation stimuli compared to the narrow condition, while expanding spatial frequency bandwidth alone had no such effect. The mixed condition produced similar responses to the ORI condition, confirming this pattern (Fig. 16, all full-field responding neurons).

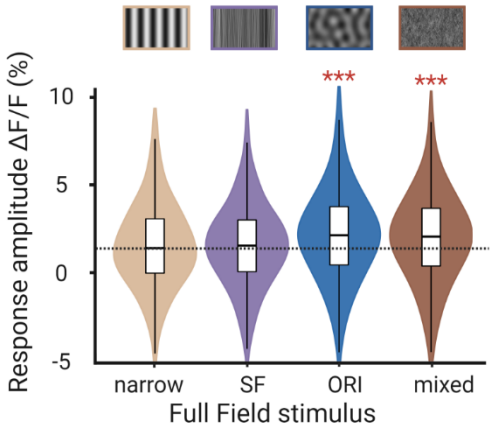


Figure 16. Neural response amplitudes were higher in the full-field mixed broadband condition.

Difference between mean response amplitude versus baseline for each full-field stimulus category. Shown are all neurons with a significant response to either of the full-field stimuli (2959 responsive neurons out of 8210 neurons in total). The black dotted line shows the median for the narrow stimulus. Response amplitudes: narrow = 3.37% ± 0.19%; SF =3.30% ± 0.18%, Wilcoxon signed rank against narrow full-field response, p = 0.98; ORI = 3.95% ± 0.20%, p = 8 x 10^{-31}; mixed = 3.90% ± 0.23%, p = 1.01 x 10^{-8}, mean ± s.e.m., n = 2959 neurons from 5 mice. For visualization only, outliers were excluded from distributions. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). The significance threshold was α = 0.0167 after Bonferroni correction for performing three tests. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

This effect was also mostly independent of the animal's behavioral state: although running generally boosted visual responses, the ORI condition still evoked consistently stronger responses than the narrow condition during both running and resting trials (Fig. 17).

For insight into the specific effect of surround modulation on single-cell responses to broadband stimuli, we focused on neurons whose receptive field center aligned with the center stimulus location and that showed a significant response to at least one of

the center stimulus conditions (one-sided Mann-Whitney U test, $p < 0.05$, see Methods). We also used pupil tracking to confirm that the mice maintained stable gaze, ensuring that receptive field positions on the screen remained consistent (Fig. 15).

For full-field stimuli, these neurons showed stronger response amplitudes in the ORI condition, with a similar upward trend observed in the mixed condition (Fig. 14d, $\Delta F/F_{narrow}$ = 3.27% ± 0.5%; $\Delta F/F_{SF}$ = 2.56% ± 0.35%, Wilcoxon signed rank against narrow, p = 0.11; $\Delta F/F_{ORI}$ = 3. 76% ± 0.49, $p = 5.9 \times 10^{-5}$; mixed = 2.98% ± 0.32, p = 0.08; mean ± s.e.m., n = 378 neurons from 5 mice). Importantly, this effect disappeared when stimuli were restricted to the receptive field center without stimulating the visual surround. In this case, neural responses were similar between the narrow and ORI conditions, and responses to the SF and mixed conditions were actually lower than those to the narrow condition. (Fig. 14e, $\Delta F/F_{narrow}$ = 4.42% ± 0.37%; $\Delta F/F_{SF}$ = 4.39% ± 0.52%, Wilcoxon signed rank against narrow, $p = 7.8 \times 10^{-3}$; $\Delta F/F_{ORI}$ = 4.56% ± 0.45%, p = 0.66; mixed = 4.47% ± 0.51%, p = 0.013; response amplitude mean ± s.e.m., n = 378 neurons from 5 mice).

These findings suggest that the increase in neural responses to broadband stimuli is not mainly due to stronger feedforward input to neurons with diverse tuning. Instead, it likely results from a reduction in surround inhibition. To assess this, we calculated a surround modulation index (SMI) for each condition, defined as the difference between the full-field and center response divided by their sum. A more negative SMI indicates stronger surround inhibition for a given neuron. We observed clear surround inhibition in the narrow stimulus condition, which remained for the broad spatial frequency bandwidth. However, SMI values were significantly less negative, indicating reduced surround inhibition for the broadband orientation and mixed conditions (Fig. 14f, SMI_{narrow} = -0.04 ± 0.01; SMI_{SF} = -0.04 ± 0.01, Wilcoxon signed rank test against narrow, p = 0.06; SMI_{ORI} = -0.02 ± 0.01, $p = 2.28 \times 10^{-7}$; SMI_{mixed} = -0.02 ± 0.01, $p = 1.42 \times 10^{-5}$, surround modulation index mean ± s.e.m., n = 378 neurons from 5 mice).

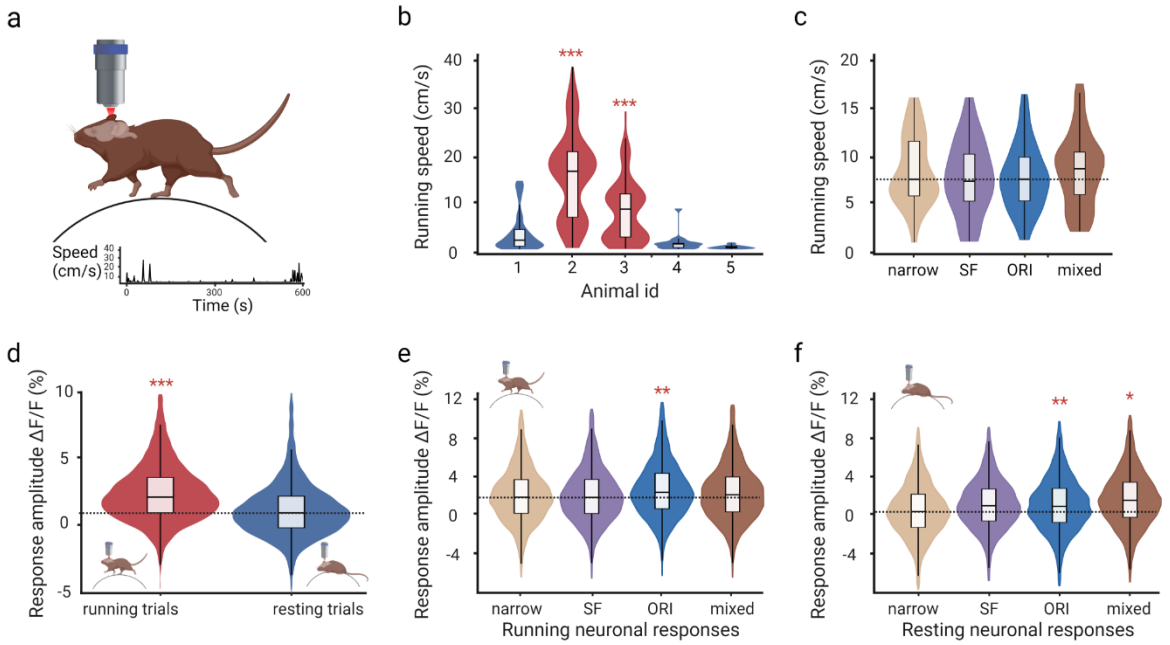


Figure 17. Broadband orientation stimuli increase neural responses during different behavioral states.

(**a**) Illustration of behavioral data acquisition during 2-photon imaging. Wheel movements were collected via a rotary encoder (bottom trace). (**b**) The average running speed varied across individual animals and was roughly split between consistent runners (animals 2 and 3) and rather passive observers (animals 1, 4, and 5, n = 600 trials per mouse, 5 mice, Wilcoxon ranked sum test against the median across all mice). (**c**) Average running speed across all presentations of each stimulus condition for all mice. There were no significant differences in running speed across conditions. Dotted line is median of the narrow stimulus average running speed (n = 55 sessions per stimulus across 5 mice, Wilcoxon ranked sum test against the other stimuli). (**d**) Running increases visual response amplitude. Shown are average response amplitudes to all visual stimuli when mice were either resting or running on the wheel. Black dotted line is the median response amplitude of resting trials. Responses to visual stimulation were significantly larger during running compared to resting trials (n = 5043 responding cells, 5 mice, LME model comparison). (**e**) Neural responses are increased for broad orientation bandwidth during running trials. Shown are the average response amplitude across all running trials for all neurons in each stimulus condition. Dotted line is median of the narrow stimulus response amplitude (n = 4541 responding cells, 5 mice, LME model testing against narrow). (**f**) Neural responses are increased for broad orientation bandwidth during resting trials. Shown are average response amplitudes across all resting trials for all neurons in each stimulus condition. Dotted line is median of the narrow stimulus response amplitude (n = 4322 responding cells, 5 mice, LME model testing against narrow). For visualization only, outliers were excluded from distributions in panels b-h. Box plots indicate the median (horizontal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). The significance threshold was adjusted after Bonferroni correction for performing the pertaining number of tests (α = 0.01, 5 tests, panel b; α = 0.0167, 3 tests, panels c, e, f). Panel a was created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

Expanding the orientation bandwidth reduces center-surround inhibition in cortical neurons, likely because the receptive field center becomes less predictable from the surround. Additionally, the lower center responses observed for the SF condition suggest that broadband spatial frequency stimuli may provide weaker feedforward input to V1 neurons compared to narrow stimuli. This reduced input could help explain the absence of increased neural responses to full-field SF stimuli.

To evaluate the role of different contributing factors, we selected surround modulation, orientation tuning, and center responses as regressor for predicting how individual neurons respond to broadband stimuli. Center responses meant to capture changes in feedforward input, therefore we calculated the difference in center stimulus responses between each broadband condition and the narrow condition. Using again an LME model with multiple regressors, we included all three factors as predictors for broadband modulation in each stimulus condition (Fig. 14g, see Methods). The model successfully predicted broadband modulation of individual neurons across all three conditions (Fig. 14h). While each factor contributed differently depending on the stimulus condition, surround modulation had the strongest and most consistent effect across all broadband stimuli. Orientation tuning significantly influenced responses to the orientation bandwidth expansion, and center response differences had a smaller but consistent effect in all cases (Fig. 14h, i).

Overall, these findings indicate that surround modulation plays a key role in shaping neural responses to broadband stimuli. The increased responsiveness to broad orientation bandwidths likely involves not just the recruitment of neurons with varied orientation tuning, but also a release from surround inhibition driven by the lower predictability of the receptive field center.

2.5 Broadband visual stimuli lead to increased neural responses in both cortical and subcortical areas.

Previous studies suggested that broadband stimuli affect neural responses differently across different HVAs (Han et al., 2022). To investigate this, we used widefield calcium imaging to measure how narrow, SF, ORI, and mixed stimuli influenced activity across visual areas. We recorded from awake, head-fixed mice positioned on a wheel and presented full-field stimuli while imaging cortical responses through the cleared, intact skull.

All imaging data was aligned to the Allen Common Coordinate Framework (Q. Wang et al., 2020) and we confirmed the locations of V1 and HVAs using visual field sign mapping (Garrett et al., 2014; Kalatsky & Stryker, 2003; Marshel et al., 2011; Zhuang et al., 2017) (Fig. 18a). Presentation of the narrow stimulus elicited strong activation in V1 and surrounding HVAs (Fig.18b, top left). As expected, responses were strongest in the left hemisphere, contralateral to the stimulus. Some ipsilateral activity was also observed, particularly in the binocular region of V1, which could be due to light reflections reaching the opposite eye.

For the analysis, we only focused on the visual areas in the contralateral hemisphere, since that is relevant to out unilaterally shown stimulus. Expanding spatial frequency and/or orientation bandwidth increased cortical activity, but this increase was more strongly confined to V1 (Fig. 18b). To confirm this, compared the response difference between ORI and narrow stimuli across V1 and HVAs (Fig. 18c). V1 showed the largest increase, with a mean difference of 0.81% ± 0.08%, supporting that broadband modulation was strongest in V1 compared to the surrounding HVAs (mean ± s.e.m.; ORI versus narrow: p = 0.0048, T = 4.34, LME model, n=8 sessions from 4 mice). Broadband orientations evoked significantly higher activity than narrow or SF in V1 in particular (Fig. 18d, mean V1 activity: $\Delta F/F_{narrow}$ = 1.90% ± 0.22%; $\Delta F/F_{SF}$ = 2.24% ± 0.14%, LME test against narrow: T = 1.73, p = 0.11; $\Delta F/F_{ORI}$ = 2.71% ± 0.25%, T = 4.34, $p = 7 \times 10^{-4}$; $\Delta F/F_{mixed}$ = 2.52% ± 0.17%, T = 2.82, p = 0.014, mean ± s.e.m., n = 8 recordings from 4 mice).

Responses to the mixed stimuli were increased but were not significantly stronger than those to the broad orientation bandwidth alone (LME test for mixed versus ORI, T = -0.77, p = 0.45). To further examine how expanding orientation and spatial frequency

bandwidth affects neuronal responses across all cortical layers, we conducted electrophysiological recordings in V1 using high-density Neuropixels probes (Jun et al., 2017) in awake mice (Fig. 18e, left). As before, the mice were head-fixed and allowed to move on a wheel, and we presented the same set of stimuli described earlier.

All stimulus conditions clearly modulated local field potentials (LFPs), with significantly stronger response modulation observed during broadband stimulation compared to the narrow condition. (Fig. 18e, right; mean LFP response difference to narrow across layers: ΔLFP_{SF} = 4.28µV ± 1.32µV, T = 3.71, p = 1.6 x 10^{-3}; ΔLFP_{ORI} = 6.44µV ± 2.82µV, T = 2.91, p = 9.5 x 10^{-3}; ΔLFP_{mixed} = 9.72µV ± 2.6µV, T = 3.78, p = 1.3 x 10^{-3}; mean ± s.e.m., LME model test against zero, n = 6 recordings from 3 mice).

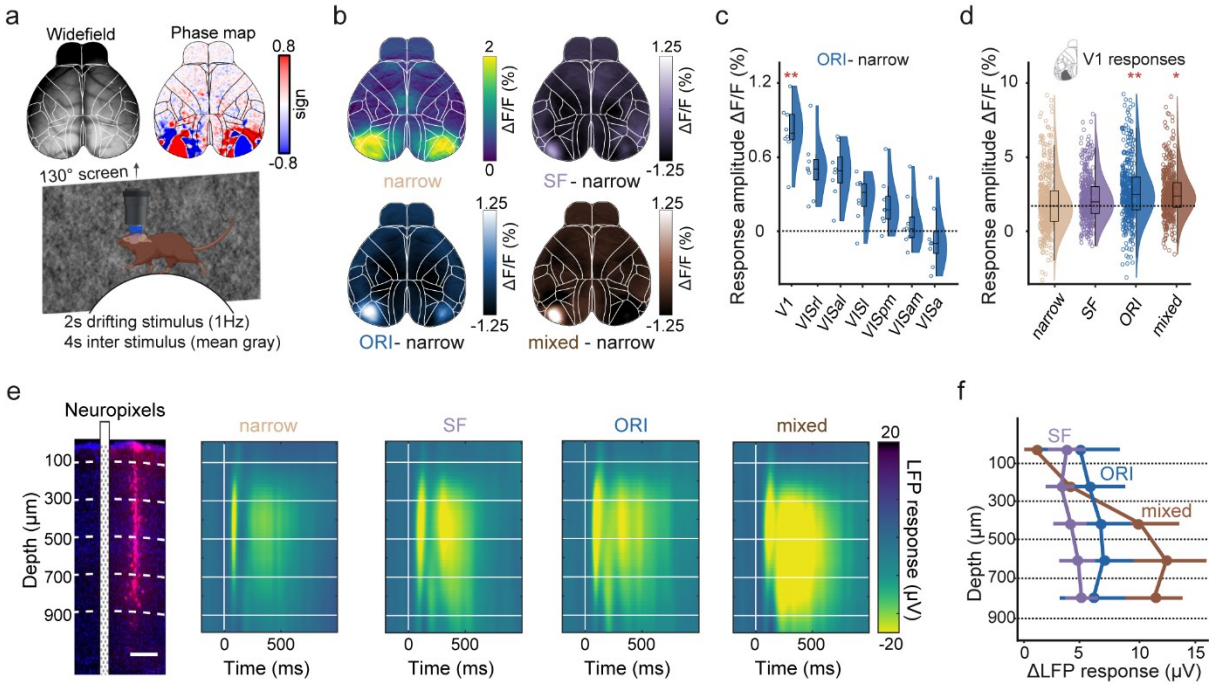


Figure 18. Broadband visual stimuli increase sensory responses in cortical and subcortical neurons.

(**a**) Overview of the widefield setup and example visual field sign map. Sign maps were used to confirm the anatomical alignment of recordings to Allen CCF (Wang, Q. et al. The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell 181, 936-953.e20 (2020)). Mice were head-fixed on a wheel and shown the same protocol as in Fig. 14 on the right-side screen. (**b**) Widefield responses averaged over the 2s stimulus period. Mean responses to the narrow stimulus condition (top left) were subtracted from mean responses to broadband stimulus conditions. Shown are average response differences for SF stimulus condition (top right), ORI stimulus condition (bottom left), and mixed stimulus condition (bottom right). (**c**) Average response differences for the ORI stimulus condition across different visual cortical areas (n = 8 sessions from 4 mice). (**d**) Individual trial response amplitudes to each stimulus integrated over the 2s stimulus period (n = 408 trials from 4 mice). (**e**) Electrophysiological recordings. Left: Example brain slice from the V1 center, showing red fluorescence from the Neuropixels probe positioning. A scheme of the probe is placed next to the fluorescent trace to visualize the probe position during the recording (scale bar: 200 µm). Different depths are marked by dashed white lines. Right: LFP responses to each stimulus condition across different cortical depths. Colors show voltage changes after stimulus

onset relative to baseline. Horizontal white lines show depth, as in the histology panel on the left. (**f**) The difference in LFP responses for each stimulus condition (SF, purple; ORI, dark blue; mixed, brown) compared to the LFP response to the narrow stimulus across depths. Error bars show mean and s.e.m. (n = 6 recordings from 3 mice). Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Stars in all panels mark significant differences from a two-sided LME model test against narrow. The significance threshold was α = 0.0167 after Bonferroni correction for performing three tests. For visualization only, outliers were excluded from distributions in panels c-d. Panel a was created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

The earliest neural responses were observed in Layer 4 (300-500 µm), followed by stronger activation in the deeper cortical layers. Consistent with our 2-photon imaging findings in superficial layers, broadband orientation stimuli produced greater response modulation than broadband spatial frequency stimuli, and this pattern held across all cortical layers (Fig. 18f). Interestingly, when both orientation and spatial frequency bandwidth were expanded, the resulting responses were markedly stronger in the deeper cortical layers, suggesting an enhanced integrative effect of broadband stimulation at those depths.

Consistent with our 2-photon imaging results, we also observed a clear increase in spiking activity in superficial layer 2/3 neurons in response to broadband orientation stimuli (Fig. S4a). To quantitatively compare spiking responses across the four stimulus conditions, we calculated the area under the receiver-operator characteristic curve (AUC). This metric provides a standardized measure of how well neural activity distinguishes between the baseline and stimulus periods, making it particularly suitable for detecting subtle responses in sparsely active neurons.

An AUC of 0.5 indicates no difference between baseline and stimulus responses (i.e., unresponsive cells), while values closer to 1 or 0 reflect consistent increases or decreases in firing, respectively. Using this measure, we found that neural responses to the ORI condition were significantly stronger than those to the narrow condition (Fig. S4a, AUC_{narrow} = 0.69 ± 0.01; AUC_{SF} = 0.68 ± 0.01, LME model test against narrow, T = -0.79, p = 0.43; AUC_{ORI} = 0.73 ± 0.01, T = 2.72, p = 7 x 10^{-3}; AUC_{mixed} = 0.70 ± 0.01, T = 0.57, p = 0.57, mean ± s.e.m., n = 53 neurons from 3 mice).

In line with our LFP findings, the mixed stimulus condition elicited stronger spiking responses in the deeper cortical layers (500-1000 µm depth, Fig. 18f) with a corresponding shift in AUC values (Fig. S4b, AUC_{narrow} = 0.70 ± 0.01; AUC_{SF} = 0.68 ± 0.01, LME model test against narrow, T = -2.77, p = 5 x 10^{-3}; AUC_{ORI} = 0.71 ± 0.01, T

= 0.94, p = 0.34; AUC_{mixed} = 0.73 ± 0.01, T = 2.96, p = 3 x 10^{-3}; mean ± s.e.m., n = 200 neurons from 3 mice).

To examine visual responses beyond the cortex, we extended our experiments to record spiking activity in the superior colliculus (SC), although we were not able to map the receptive fields of SC neurons (Fig. S5). Interestingly, SC neurons showed much stronger responses to the mixed stimulus condition and, to a lesser extent, to broadband orientation and spatial frequency stimuli, compared to the narrow condition. (Fig. S5d, AUC_{narrow} = 0.68 ± 0.02; AUC_{SF} = 0.68 ± 0.02, LME model test against narrow, T = 0.11, p = 0.91; AUC_{ORI} = 0.70 ± 0.02, T = 1.10, p = 0.27; AUC_{mixed} = 0.82 ± 0.02, T = 5.44, p = 4.46 x 10^{-6}, mean ± s.e.m., n = 46 neurons from 3 mice).

Taken together, these findings show that the increased cortical activation caused by broadband visual stimuli is primarily confined to V1. In the cortex, superficial layers show a clear preference for broadband orientation stimuli, while deeper layers respond strongly to mixed stimuli that combine broadband orientation and spatial frequency. Interestingly, this pattern is also seen in the subcortical superior colliculus, which receives direct input from the retina as well as from the deeper layers of V1.

2.6 Broader orientation bandwidth leads to an enhanced ability to discriminate between visual stimuli in the neural network.

We next asked whether expanding orientation bandwidth not only boosts neural responsiveness but also enhances the neural encoding of sensory information, specifically, whether it improves the ability of neurons to discriminate between different central features. To test this, we presented stimuli differing in either central orientation (0° vs. 90°) or central spatial frequency (0.04 cpd vs. 0.16 cpd), while varying the bandwidth of the non-tested feature. (Fig. 19a, e). This approach allowed us to isolate the effect of bandwidth expansion without introducing ambiguity. Spatial frequency bandwidth was varied during orientation discrimination, and orientation bandwidth was varied during spatial frequency discrimination. In line with previous findings, increasing orientation bandwidth (with a fixed central spatial frequency of 0.16 cpd) significantly increased the number of visually responsive by a factor of 1.36 ± 0.14 fold compared to the narrow condition (Fig. 19b, T = 2.75, $p = 9 \times 10^{-3}$; LME test against 0; percentage responsive neurons over total neurons per session: narrow range (5°): 11% ± 1%, mid range (25°): 15% ± 2%, broad range (45°): 21% ± 2%, mean ± s.e.m, n = 18 sessions from 9 mice). This increase was smaller than that observed for a central frequency of 0.04 cpd, likely because high-frequency stimuli are less effective at driving V1 neurons (LME test, T = 3.91, $p = 4.1 \times 10^{-5}$), (Fig. 14e).

Alongside the increase in the number of responsive neurons, response amplitudes also rose, indicating that orientation bandwidth expansion enhances V1 responses regardless of central spatial frequency (Fig. 19c; orientation modulation index: ORI $RMI_{0.04cpd}$ = 0.17 ± 0.01, LME test against zero, T = 14.3, $p = 7.27 \times 10^{-45}$; ORI $RMI_{0.16cpd}$ = 0.06 ± 0.01, T = 4.3, $p = 1.53 \times 10^{-5'}$; mean ± s.e.m., n = 1345 neurons from 18 recordings from 9 mice). This broader activation likely enhances the representation of other features, like spatial frequency, by increasing the number and strength of responding neurons.

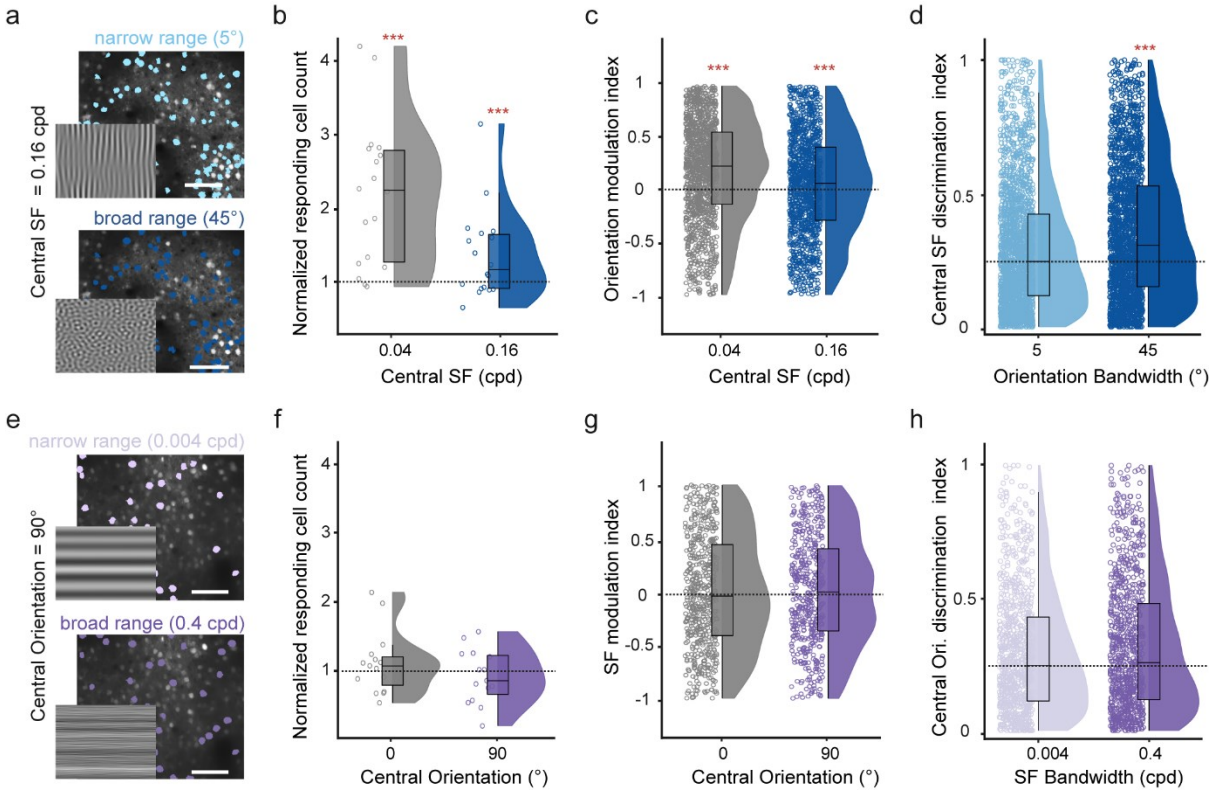


Figure 19. Expanding orientation bandwidth enhances stimulus discriminability in V1 neurons.

(**a**) Example field of view of an experiment with the central spatial frequency of 0.16 cpd and narrow or broad orientation bandwidth. Stimuli are shown at the bottom left of the plane, scale bar: 100 µm. (**b**) Number of responsive cells to broad orientation bandwidth (45°) with either 0.04 cpd or 0.16 cpd central spatial frequency, normalized to the respective number of cells responding to the narrow condition (5°). Data for 0.04 cpd are the same as in Fig. 7 and shown here for reference in gray. (**c**) Orientation modulation index, calculated as mean broad minus narrow orientation bandwidth responses, divided by their sum. Shown are all positively narrow and broad responding neurons for each central spatial frequency (n = 1333 responsive neurons for central SF = 0.04 cpd in gray, similar to data in Fig. 8c, n = 1345 responsive neurons for central SF = 0.16 cpd in blue. n = 4892 neurons in total across 18 sessions from 9 mice). (**d**) Discriminability between neural responses to stimuli with a central spatial frequency of either 0.04 cpd or 0.16 cpd. Shown is the discrimination index for narrow and broad orientation bandwidth in light and dark blue, respectively (n = 2096 neurons from 18 sessions). Shown are the results for all cells that responded to any of the presented stimuli. The horizontal black dotted line shows the median discriminability for the narrow-band. (**e**) Example field of view of an experiment with 90° central orientation and narrow or broad frequency bandwidth, scale bar: 100 µm. (**f-h**) Same as panels **(b-d)** but for narrow and broad frequency bandwidth and 0° and 90° central orientation. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Stars in all panels mark significant (Bonferroni correction for two tests, α = 0.025) differences from a two-sided LME test against 1 (panels c, g) or the narrow bandwidth (panels b, d, f, h). For visualization only, outliers were excluded from distributions in panels b-d, f-h. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

Crucially, neuronal responses to broadband stimuli with different central orientations or spatial frequencies were not highly correlated, suggesting that neurons respond selectively based on the central feature, even under broadband stimulation (Fig. 20).

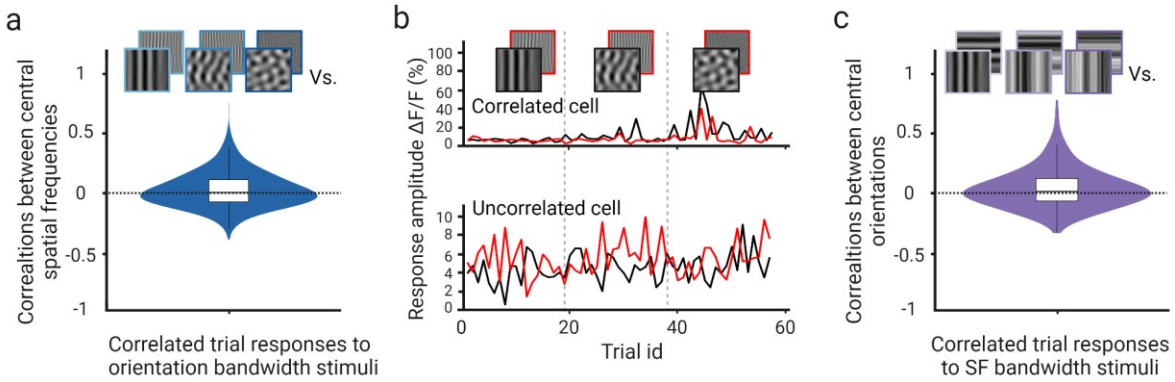


Figure 20. Bandwidth selectivity varies depending on the central spatial frequency or orientation.

(**a**) Neuronal response correlations were analyzed for stimuli with different orientation bandwidths and either low (0.04 cpd) or high (0.16 cpd) central spatial frequency (SF). Stimulus examples shown on top. Responses from neurons significantly activated by any bandwidth (19 trials per bandwidth, 57 responses in total) were correlated across SF conditions. No significant correlation was found (r = 0.02 ± 0.003, mean ± s.e.m., n = 2097 neurons, Wilcoxon test, p = 0.073), indicating no systematic preference for a specific bandwidth independent of SF. Additionally, bandwidth preference showed little overlap across SFs: 9% of narrow-preferring neurons at 0.04 cpd remained narrow-preferring at 0.16 cpd, 5% for mid-preferring, and 13% for broad-preferring. Note that the overlap of bandwidth-selective neurons across the central SFs was not a result of the correlation analysis but is mentioned here for completeness. (**b**) Response amplitudes to all stimuli for two example cells from a), stimuli shown above. Top cell showed a high response correlation and preferentially responded to broad orientation bandwidth stimuli with either central SF. Dotted lines mark the last trial for each bandwidth condition. Red lines - high central SF responses, black lines - low central SF responses. Bottom example shows a de-correlated neuron response vector. (**c**) Same as a) for correlations to stimuli with different SF bandwidths and a central orientation of either 0° or 90°. Lack of significant difference from 0 suggests that SF bandwidth selectivity of neural responses is not independent of the central orientation (Pearson correlations: r_{value} = 0.04 ± 0.005, mean ± s.e.m., n = 888 neurons, Wilcoxon signed-rank test against 0, p = 0.069). Again, little overlap between the SF bandwidth selective neurons across the two central orientations: 8% of narrow SF bandwidth preferring neurons at 0° central orientation were also narrow SF bandwidth preferring at 90° central orientation, 3% of mid preferring at 0° were also mid preferring at 90° and 6% of broad preferring at 0° were also broad preferring at 90°. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

To directly test whether broader orientation bandwidth improves the neural discriminability of spatial frequencies, we calculated a discrimination index (AUC_{abs}) as the absolute AUC between neural responses to stimuli with a central spatial frequency of either 0.04 cpd or 0.16 cpd (see also Fig. 23 in the methods section). AUC_{abs} was normalized between 0 and 1. A higher value means enhanced ability to discriminate between central spatial frequencies. We computed the discriminability for either narrow or broad orientation bandwidth stimuli (Fig. 19d, $AUC_{abs,\,5°}$ = 0.30 ± 0.01; $AUC_{abs,\,45°}$ = 0.36 ± 0.01, LME model test against narrow condition, T = 8.13, p = 5.44 x 10^{-16}; mean ± s.e.m., n = 2096 neurons from 18 sessions, 9 mice). The significantly higher AUC with broader orientation bandwidth indicates enhanced neural discriminability of spatial frequency, suggesting that orientation bandwidth expansion improves the representation of other stimulus features.

In contrast, increasing spatial frequency bandwidth didn't improve responses to stimuli with central orientations of 0° or 90° (Fig. 19f, number of visually-responsive neurons compared to the narrow condition: broad SF bandwidth (90° central orientation) = 1.09 ± 0.21, T = 0.42, p = 0.67, LME model test against zero, mean ± s.e.m; Fig. 19g, spatial frequency modulation index: SF $RMI_{0°}$ = -0.004 ± 0.02, T = -0.17, p = 0.86; SF $RMI_{90°}$ = 0.014 ± 0.02, T = 0.6, p = 0.53, mean ± s.e.m., n = 16 recordings from 9 mice).

Consequently, we observed no significant change in the neural discriminability of stimuli with different central orientations when spatial frequency bandwidth was expanded (Fig. 19h, orientation discrimination index: $AUC_{abs,\ 0.004cpd}$ = 0.30 ± 0.01; $AUC_{abs,\ 0.4cpd}$ = 0.32 ± 0.01, LME model test against narrow condition, T = 1.8, p = 0.06, mean ± s.e.m., n = 888 neurons from 9 mice).

In summary, expanding orientation bandwidth increases the number and strength of neural responses and significantly improves the ability of V1 neurons to distinguish between different central spatial frequencies. In contrast, expanding spatial frequency bandwidth does not improve orientation discriminability, highlighting a key asymmetry in how V1 neurons encode complex visual features.

2.7 Broader orientation bandwidth leads to an enhanced ability to discriminate between visual stimuli in the mouse.

Our results show that expanding orientation bandwidth enhances neural responses in V1 and improves the discriminability of broadband visual stimuli at the neuronal level. To determine whether this also leads to improved visual perception, we trained mice to perform a visual discrimination task while freely moving in a custom-built touchscreen chamber (Wiesbrock et al., 2022) (Fig. 21a). In this setup, mice were trained to touch one of two motion cloud stimuli on a touch-sensitive screen and received a reward for correctly identifying the target stimulus.

We tested two groups of mice (6 per group) trained to discriminate either central spatial frequency (0.16 cpd target vs. 0.04 cpd non-target; Fig. 21, left) or central orientation (0° target vs. 90° non-target; Fig. 21b, right). For both tasks, the bandwidth of the non-tested parameter was varied to examine its effect on performance, i.e., spatial frequency bandwidth was varied during orientation discrimination, and orientation bandwidth during spatial frequency discrimination (Fig. 21b). Both groups learned their tasks reliably and reached comparable expert performance.

Notably, expanding orientation bandwidth improved spatial frequency discrimination: mice performed better when spatial frequency differences were judged under broadband orientation conditions (Fig. 21c, percentage of correct trials: narrow (5°) = 77.0% ± 1.93%; mid range (25°) = 77.63% ± 1.62%, Wilcoxon signed rank against narrow condition: p = 0.92; broad (45°) = 82.77% ± 1.41%, $p = 4.9 \times 10^{-3}$, mean ± s.e.m., n = 46 sessions from 6 mice). This indicates that broader orientation content not only enhances neural discriminability but also improves perceptual sensitivity to spatial frequency.

In contrast, expanding spatial frequency bandwidth did not improve orientation discrimination. In fact, there was a trend toward decreased performance with broader spatial frequency bandwidth (Fig. 21d, percentage of correct trials: narrow (0.004 cpd) = 72.53% ± 2.85%; mid range (0.04 cpd) = 69.18% ± 2.12%, Wilcoxon signed rank test against narrow percentage of correct trials, p = 0.79; broad (0.4 cpd): 65.34% ± 3.60%, p = 0.29, mean ± s.e.m., n = 24, 16 and 18 sessions for each condition from 6 mice).

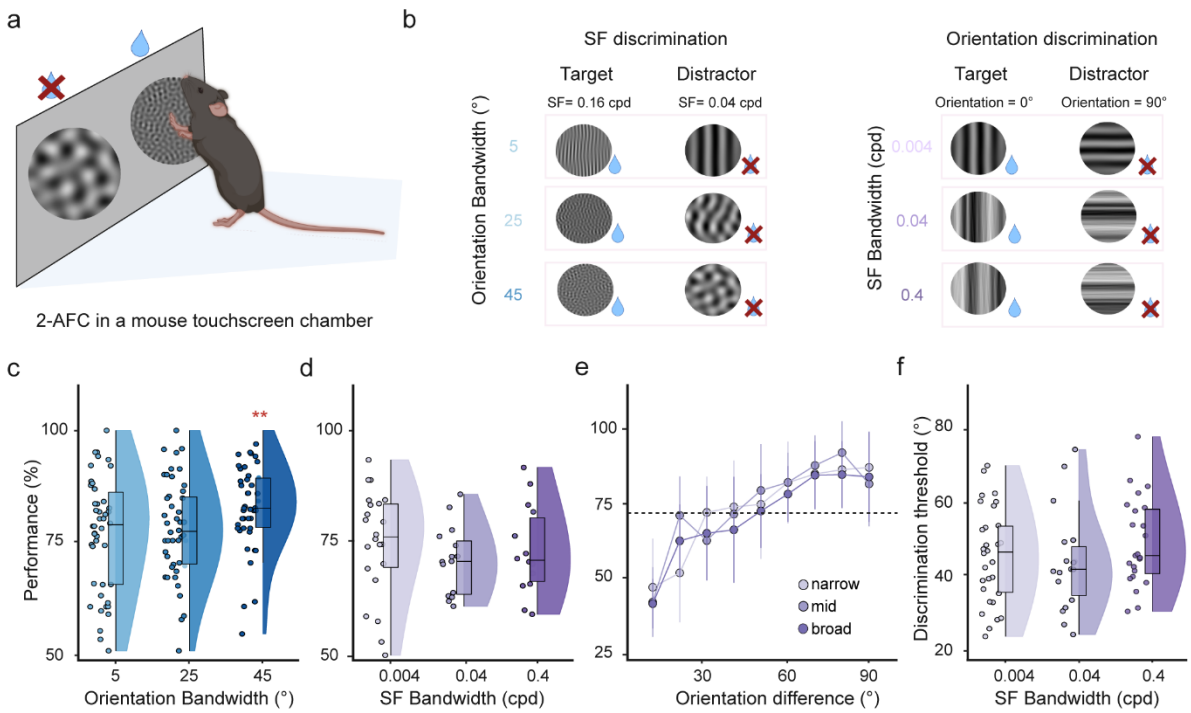


Figure 21. Increasing orientation bandwidth leads to improved visual perception.

(**a**) Illustration of the visual discrimination task. The mouse initialized the trial by triggering the lick detection, followed by a 500-ms long ITI with a gray screen, followed by a visual cue (white cross) for 700 ms. Two stimuli were then shown next to each other, and the mouse had to touch one stimulus to report a choice. Mice then received a water reward at a central lick spout after touching the target stimulus. Triggering the lick detection (even if no reward was given) initiated the next trial. (**b**) Example stimuli for the spatial frequency discrimination task with different orientation bandwidths (left) and the orientation discrimination task with different spatial frequency bandwidths (right). (**c**) Spatial frequency discrimination performance for different orientation bandwidths. Spatial frequency discrimination performance was significantly higher for larger orientation bandwidth (n = 46 sessions from 6 mice). (**d**) Same as in panel c) but for the orientation discrimination performance with different spatial frequency bandwidth stimuli. (n = 25,17,19 sessions respectively.) (**e**) Psychometric curves for orientation discrimination performance at different target-distractor orientation differences. Colors show psychometric curves for different spatial frequency bandwidths. Expanding the spatial frequency bandwidth did not affect the maximal discrimination performance or discrimination thresholds. The horizontal dashed line shows the 72.7% discrimination threshold. Error bars are centered at the colored circles and show standard deviation (n = 30,18 and 24 sessions, respectively). (**f**) Discrimination thresholds for the three spatial frequency bandwidths showed no significant differences (n = 30, 18 and 24 sessions respectively). Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Stars in panels c, d, and f show the significance of a two-sided Wilcoxon signed rank test against the narrow condition. The significance threshold was α = 0.025 after Bonferroni correction for performing 2 tests. Panel a was created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

These results strongly suggest that enhanced V1 discriminability for broadband orientation stimuli translates into better perceptual discrimination. The lack of change with spatial frequency bandwidth might reflect a ceiling effect, where differences in perception become more noticeable only near the perceptual threshold. To test this, we used a staircase procedure that dynamically adjusted orientation differences between target and non-target stimuli to measure individual orientation discrimination thresholds (Fig. 21e). Thresholds and psychometric functions were similar across

spatial frequency bandwidths, again showing no significant influence of spatial frequency bandwidth on orientation perception (Fig. 21f, discrimination threshold at 72.7% orientation discrimination threshold from staircase procedure: narrow (0.004 cpd) = 46.54° ± 2.33°; mid range (0.04 cpd) = 49.93° ± 2.53°, Wilcoxon signed rank test against narrow, p = 0.61; broad range (0.4 cpd) = 49.34° ± 2.24°, p = 0.28, mean ± s.e.m, n = 18, 23 and 30 sessions respectively).

In summary, these findings demonstrate that broadband orientation stimuli, not broadband spatial frequency stimuli, increase neural responsiveness, enhance stimulus discriminability in V1, and improve perceptual performance.

3. DISCUSSION

This study was focused on neuronal processing of nature inspired, feature-enriched visual stimuli and their impact on the visual perception in mice. Random phase textures of differing orientation or spatial frequency bandwidths were shown to awake, passively viewing mice to get access to their neuronal responses to the different feature richness levels, or the stimuli were integrated into a discrimination task to get a read-out of the perceptual impact of these neuronal responses. We used different perspectives of looking at the neuronal responses of their visual cortex neurons. First, looking at the calcium-inferred population activity in V1, layer 2/3, via 2-photon imaging, we saw that broadening the orientation bandwidth of the stimuli increased neuronal recruitment as well as neuronal responses (Figs. 7, 8). Then, we tested if this increase in responses was merely due to recruitment of diversely orientation-tuned V1 neurons. For this, a Gabor filter bank model simulating V1 responses was built (Figs. S1-S2). Orientation tuning alone could not explain the effect. An orientation tuned divisive surround suppression component needed to be added to account for the behaviour of real V1 neurons. The impact of surround suppression was then experimentally confirmed by additional 2-photon imaging (Figs. 12, 14). Then, we looked at the cortex-wide activity to localize potential focal areas or HVA contributors. Though V1 turned out to be the area with the strongest orientation bandwidth enrichment-driven activity increase, there were few lateral HVAs impacted as well. Also, we looked at the spiking activity of neurons at different depths via Neuropixel probes, where deeper cortical areas stood out with their strong preferences for a stimulus that is rich in both orientation and spatial frequency bandwidths, hence highly unpredictable (Fig. 18). The predictability aspect of the stimuli is important because it also defines the mode of representation in the neuronal population. Due to decreased predictability, increased response saliency, and increased separation in population representation, richer orientation bandwidth stimuli of different central spatial frequencies are also better discriminated by the V1 neuronal population (Fig. 19). Additionally, this improved discrimination capability also manifests at a behavioural level, with mice becoming better at discriminating different central spatial frequencies upon orientation bandwidth enrichment (Fig. 21). These conclusions support a predictability and context-dependent surround modulation of visual information at the

cortical level. Scenes that are highly unpredictable and rich in orientation features drive a reduction in surround inhibition, which allows stronger recruitment of V1 neurons, increased neuronal response strength, a better representation of the visual information, and improved visual perception.

3.1 The representation of orientation bandwidth enrichment.

The central finding of this study was that broadening the orientation bandwidth of random phase textures increased the responses of neurons in the mouse V1 across all layers, and this increase was present in several other lateral visual areas, as well as several cortical layers in V1 and spanning as far as the optic SC. The following subsections dissect the neuronal representation of orientation bandwidth enrichment and the impact this enrichment has on perception.

3.1.1 Mouse V1 network dynamically adjusts its orientation tuning selectivity in response to higher-order visual features.

Our observation that V1 recruits more neurons in response to orientation bandwidth enrichment adds to previous studies that have shown an increase in V1 responses for broadband orientation frequency and spatiotemporal frequency noise stimuli (Goris et al., 2015). Knowing that the majority of V1 neurons are sharply orientation tuned and each one of them prefers a certain range of orientations, the simplest assumption would be that the increase in recruitment could be explained merely by the activation of additional neurons, which have a different tuning curve from the neurons which respond to the narrow orientation band stimuli. To test if the matter is that trivial, we built a model composed of several Gabor filters with similar size tuning, spatial frequency tuning, and orientation tuning as the mouse V1 population (Marshel et al., 2011; Roth et al., 2012). The output from this model tends to overshoot and doesn't match the real neuronal activity. Looking at the low spatial auto-correlations of broadband orientation stimuli, it stands out that they are spatially unpredictable and would require more processing resources in comparison to the narrow orientation band stimuli (with a higher spatial auto-correlation). Neurons in the HVAs have larger

receptive fields and could influence the modulation of responses to redundant visual information (Andermann et al., 2011; Marshel et al., 2011). We added a surround suppression component to the model, and that helped the model output to better match the mouse neuronal population responses. We also confirmed this experimentally and built an LME model to break down the contribution of different components into the broadband orientation response enrichment and saw that orientation tuning, as well as feedforward input, have a significant impact. However, surround modulation is the strongest contributor (see Fig. 14).

These considerations lend support to the argument that rather than operating as a mere bank of linear filters, the V1 network selectively adjusts its tuning based on the higher-order features of the visual input and the spatial arrangement of the orientation information in the scene. This signal optimization strategy is reminiscent of predictive coding principles, whereby expected, dominant or redundant features are attenuated.

More mouse V1 neurons are tuned to cardinal orientations (Roth et al., 2012), however that has been confirmed via the use of one-dimensional stimuli such as square-wave gratings, which doesn't allow for a perspective into the population employment mode in complex stimulus context which is more commonly found natural scenes. The distribution of orientation tuning in the mouse V1 neuronal population reflects natural statistics, but that doesn't mean that the activation pattern needs to employ a selection of neurons that reflect that distribution of tuning. It could instead make use of this representation through a different strategy. Looking at the neuronal responses to broadband stimuli sorted by tuning (Fig. 12 c), one notices that the cardinal orientation over-representation from the narrow band stimuli is lost, despite the cardinal orientation still being the central orientation in the stimulus (Fig. 7b). This exposes how in the narrow band stimuli the dominant cardinal stimuli enhance rather than attenuate tuning selectivity. A possible explanation for this could be that 45 degree tuned neurons have broader tuning width and are therefore more strongly activated by the broad stimuli than 0 degrees tuned cells.

Intuitively, one would have to wonder how generalizable this strategy is. Mice don't have structured columns of orientation tuned cells clustered together like primates do. Instead they have a salt and pepper pattern of orientation tuned cell spatial clustering.

Thus, their network tuning is likely to be more malleable, more dependent on visual contexts such as salient edge orientations, and the dynamic modulation observed here might reflect a compensation for the lack of long-range structured connectivity.

The efficiency of re-weighing the network tuning sharpness for different orientation contexts becomes more apparent while considering the population code. A population code that is shaped by the global structure of the input, rather than the individual independently tuned neurons, allows for rapid adjustments to diverse sensory inputs, without the requirement of synaptic plasticity. This aligns with recently observed, rapid state-dependent reshaping of population codes in the sensory cortex (Musall et al., 2019; Stringer et al., 2019).

Overall, our findings further strengthen the idea that orientation tuning in mouse V1 is dynamic and context sensitive, and that this is especially revealed by natural visual scene statistics. This context sensitivity may arise from a combination of intrinsic neuronal properties, feedforward input normalization and top-down modulation, which are further explored in the following subsections.

3.1.2 Surround suppression regime depends on the spatial predictability.

Surround suppression has mostly been studied in paradigms where a sine or square wave grating stimulus extends beyond the preferred receptive field of a neuron, driving a reduced response compared to when the stimulus was being contained within that receptive field (Adesnik et al., 2012; K. Guo et al., 2005; A. J. Keller, Roth, et al., 2020; Pecka et al., 2014). Thus, this neuron is suppressed by the simultaneous stimulation of the surround as well as the center of the receptive field. The narrow band orientation stimulus showed a similar effect, whereas for the broadband orientation stimulus this suppression effect was reduced drastically.

Surround suppression has been shown to depend on orientation similarity between the center and the surround, and it diminishes with increased dissimilarity between the two areas (Kirchberger et al., 2021; Self et al., 2014). Something that sets these bandwidth stimuli apart is the amount of spatial auto-correlation (Portilla & Simoncelli, 2000), which reduces a lot with the orientation bandwidth expansion. This translates

as increased dissimilarity between the center and the surround and as a decrease in stimulus center predictability based on the surround information. Thus, this change in surround inhibition after a reduction in center stimulus predictability, aligns well with the expectations of the predictive coding theory (Coen-Cagli et al., 2015). Therein, for resource conservation purposes, the responses to unpredictable stimuli are prioritized over responses to predictable stimuli.

Two primary cortical mechanisms are thought to contribute to surround suppression: lateral inhibition from co-oriented horizontal connections within V1 and top-down feedback from higher visual areas. Similarly, the diverse cortical responses to several superimposed grating stimuli could also be explained by V1 horizontal interactions (Muir et al., 2015, 2017; Palagina et al., 2017) or by thalamocortical feedforward mechanisms (Ruderman & Bialek, 1994).

V1 neurons have a secondary receptive field in the surround with different visual tuning properties than the center receptive field (A. J. Keller, Roth, et al., 2020). Silencing higher visual areas has shown to reduce surround suppression in V1 (A. J. Keller, Roth, et al., 2020). Anaesthesia also has a similar effect (Adesnik et al., 2012; A. J. Keller, Roth, et al., 2020), and additionally it has been shown to have a higher impact on HVA responses (A. J. Keller, Roth, et al., 2020). Furthermore, other rich stimuli such as superimposed gratings, which also increase responses in awake monkeys and mice (Barbera et al., 2022; Guan et al., 2020; Han et al., 2022; Juavinett & Callaway, 2015; Muir et al., 2015, 2017; Palagina et al., 2017; Ruderman & Bialek, 1994), have been shown to drive lower responses in anesthetized cats (Busse et al., 2009; DeAngelis et al., 1992; T. C. B. Freeman et al., 2002; Ladret et al., 2023; Ringach, 2019) and monkeys (Goris et al., 2015). V1 center-surround interactions in awake monkeys are sensitive to higher-order structures such as image contours (K. Guo et al., 2005), but they lose that sensitivity the when the Fourier spectrum phases of the image surround are randomized (Coen-Cagli et al., 2015).

Together, these findings suggest that feedback signals contribute strongly to the release from surround suppression in response to the increased unpredictability of the spatial arrangement of orientations in the broadband stimuli. From a computational perspective, surround suppression is a useful strategy that the visual system employs to refine the population code by lowering the responses to spatially uniform and

redundant spatial structures (Haider et al., 2010; Pecka et al., 2014; Vinje & Gallant, 2000, 2002). This improves the signal to noise ratio for more relevant information like edges and boundaries and helps prioritize novel information and stabilize representations by employing less neurons (Haider et al., 2010). It has not yet been formalized how the feedback signal achieves this release, but a disinhibition circuit of interneurons such as for example VIP inhibiting SOM which in turn release pyramidal neurons, is a potential candidate.

3.1.3 Broadband orientation stimuli drive strong responses in several lateral HVAs, but the strongest effect is in the granular layers of V1.

Identifying local mechanisms of naturalistic stimulus encoding in the first cortical processing step (V1) is a crucial first step in understanding the visual perception of nature. It is, however, important to place these mechanisms within the general cortical context by looking at the general input-output hierarchy in V1 and also in the global activity of the visual cortex.

HVAs, being functionally intended for more complex features than V1, certainly show their own strategies and preferences for processing visual stimuli. They have been shown to have different spatial frequency or speed tuning than V1, and generally stronger orientation tuning (Han et al., 2022; Roth et al., 2012). Also, previous studies have looked at the effects of broadband stimuli in different HVAs and noticed that there are area-specific specializations (Han et al., 2022). In line with this, several areas are more tuned to natural textures than scrambled images or simple gratings (J. Freeman et al., 2013; Yu et al., 2022).

With these considerations in mind, we sought to identify cortical areas of increased broadband tuning by looking at the global cortical activity via widefield imaging. And although we did observe a lateralized focal area of broadband orientation preference, broadband response modulation was strongest in V1. This points out the crucial role of this area in orientation rich stimulus processing in the mouse visual system. Interestingly, we also noticed that a stimulus richer in both orientation and spatial frequency information also shows a strong response modulation in V1.

At the same time, we wanted to have an overview on how the different V1 input-output hierarchy levels contribute to this broadband response modulation. For this, we did electrophysiological measurements throughout V1 to look at layer specific preferences, as well as at the optical SC for insight into the feedforward input. Moreover, SC receives most of its inputs directly from the retina and helps reveal early stage modulations (from the retina or locally in the SC) to broadband stimuli.

These measurements revealed that orientation broadband response modulation happens throughout V1 layers and even goes as far as optical SC. Interestingly, they also showed that a stimulus richer in both orientation and spatial frequency information can further increase neural responses at deeper layers. This shows that the weak impact of spatial frequency bandwidth enrichment is specific to the supragranular V1 layers. Furthermore, the mixed stimulus response modulation in the widefield and deeper LFP could be explained by dendritic signals of layer 5 neurons, which also contribute to widefield signals (Peters et al., 2021). Whereas the increased preference in the SC for the mixed stimulus could suggest that less-predictable and more salient stimuli recruit SC circuits more strongly than the visual cortex.

Different layers have been associated with different types of saliency processing. For example, while layer 5 covers the representation of the internal predictions (Harris & Mrsic-Flogel, 2013; G. B. Keller & Mrsic-Flogel, 2018), superficial V1 layers are more implicated in the sequence prediction errors (Gallimore et al., 2023), visuo-motor mismatch (Fiser et al., 2016; G. B. Keller et al., 2012; G. B. Keller & Mrsic-Flogel, 2018; Zmarz & Keller, 2016) and center-surround prediction errors (A. J. Keller, Dipoppa, et al., 2020; A. J. Keller, Roth, et al., 2020; Kirchberger et al., 2021). V1 neurons have a secondary receptive field in the surround with different visual tuning properties than the center receptive field (A. J. Keller, Roth, et al., 2020). Which, placed together in context with the tuning of center and surround interaction to the characteristics of natural scenes (Fu et al., 2023), could explain the broadband response layer-dependence.

3.1.4 How far can these effects be generalized?

Placing our results in context of their broader field and recent similar studies two important questions arise:

(1) How come natural movies sparsen neuronal responses while our naturalistic broadband orientation stimuli drive widespread and strong neuronal responses?
(2) How come cats show widespread population responses to orientation stimulus enrichment but their response amplitudes are gain normalized (Ladret et al., 2023), whereas mouse V1 population does not show this gain normalization?

Possible explanations for question (1) could lie in certain differences between a natural movie and our random phase stimulus, such as higher order structure and spatio-temporal correlations. For example, a natural movie might not always be temporally coherent but often contain static edges, sharp edges, object boundaries and slow texture transitions that drive stronger local feature selectivity due to lateral and recurrent circuits. Natural movies contain broad orientations indeed, but these orientations are not homogenously spread across the scene, not presented in a homogenous temporal flow and the broadband orientation regions are often static and localized, so neurons may respond to individual components selectively (e.g. based on expectation and deviations thereof) instead of the whole stimulus. The broadband orientation motion cloud stimuli, on the other hand, have a continuously smooth motion signal and a strong local motion energy at many different orientations, which can drive a synchronized V1 population response. This adds to the idea that motion continuity is an important unifying drive in V1 (Marques et al., 2018), especially in mice where motion is very central to their visual processing of threats, etc. Furthermore, the sparsity in responses is also related to the finding that spatial correlations of a natural scene coordinate the activation of specific subsets of neurons in the mouse V1 for increased reliability of encoding (Rikhye & Sur, 2015).

While for both cats and mice, natural images drive sparse but strong responses due to tuning to edges and structured input, broadband motion clouds show widespread activity but at different activation strength levels for these different animals. The most obvious argument to address question (2) is the cortical organization between these two animals. Cats have a columnar pinwheel organization, whereas mice the salt and

pepper distribution of feature-tuned neurons. Based on this organization, the local competition to orientation responses is stronger for cats, therefore more neurons can still be recruited while their responses are diminished. Furthermore, cats have a higher visual acuity and rely on finer spatial scales, strong edges and contrast structures, and their V1 neurons have smaller receptive fields and a sharper orientation tuning. Whereas mouse V1 neurons have larger receptive fields and are more broadly tuned in comparison, which means they integrate over more orientations and across space.

3.1.5 Broadband orientation stimuli improve visual perception.

Broadband orientation stimuli also improved spatial frequency discrimination. A driving force for this could be the refinement in population code that allows for better separation of different representations in the network. This even translates to CNNs, whereby, adding a surround suppression refinement has shown to improve their performance and training speed in visual tasks (Hasani et al., 2019).

Furthermore, the decrease in spatial predictability might also help improve the sensory perception due to increased stimulus saliency, allowing patches of the stimulus to stand out compared to the rest of the surround (Bergen & Julesz, 1983; A. J. Keller, Dipoppa, et al., 2020; Kirchberger et al., 2021; Lamme, 1995; Treisman & Gelade, 1980). This argument becomes even more relevant when looking at the strong preference for the highly unpredictable, mixed stimulus in the SC, which is shown to encode a saliency map of visual scenes in primates (White et al., 2017).

3.2 The representation of spatial frequency bandwidth enrichment.

In contrast to the orientation, broadening the spatial frequency bandwidth of the stimulus did not have as much of an impact on the supragranular V1 neuronal responses or on the mouse visual perception. It did, however, have an impact in the deeper V1 layer and in the optical SC. The following subsections dissect the neuronal representation of spatial frequency bandwidth enrichment and the impact this enrichment has on perception.

3.2.1 Mouse V1 maintains stable activity despite richer spatial frequencies.

Broadening the spatial frequency bandwidth of our random phase textures did not show a change in the neuronal recruitment in the mouse V1 L2/3 population. These neurons have been previously characterized for their spatial frequency tuning (Girshick et al., 2011; Marshel et al., 2011; Roth et al., 2012), therefore we checked if a broad tuning to this feature, resulting in activation of the same neurons by different spatial frequency bandwidths (Roth et al., 2012), was the reason for the lack of increase in neuronal recruitment. However, within our tested ranges of orientation and spatial frequency bandwidths, the neurons had comparable tuning widths. Furthermore, the population of cells that unselectively responded to orientation bandwidth stimuli was larger than the population of cells that unselectively responded to spatial frequency bandwidth stimuli. This suggests that the difference in bandwidth enrichment between these two features is not explained by the tuning properties of the neurons. Beside a very small population of stimulus-responsive neurons that respond to all spatial frequency bandwidths, an even smaller population preferred the midrange spatial frequencies, which are in between the narrow and broad and contain elements from both. This means that broader spatial frequency bandwidths activate new subpopulations of differently tuned neurons, while releasing the narrow-tuned neurons, thus, maintaining an overall status quo of the network activity.

Furthermore, a broad spatial frequency stimulus contains more of the dense structures, which the brain has evolved to decorrelate and represent sparingly. In this context, a representation that maintains overall resources at a stable level while the redundant information increases, aligns well with efficient coding strategies. This might help enhance detection of relevant frequencies.

This reduction in representation can already be seen in the reduced center responses (which represent feedforward information) to the SF stimulus compared to the narrow stimulus. Also, the spatial auto-correlations of the broadband spatial frequency stimulus revealed only a small decrease in stimulus predictability compared to the broadband orientation stimulus, making thus surround suppression more impactful on

the neuronal responses to this stimulus. This difference in stimulus predictability could also explain the difference in response enrichment between the two features.

3.2.2 Broadband spatial frequency stimuli have a subtle effect in V1.

HVAs have been shown to have different spatial frequency or speed tuning than V1 (Han et al., 2022; Roth et al., 2012) and specialized area streams for complex feature encoding have been revealed (J. Freeman et al., 2013; Han et al., 2022; Yu et al., 2022). With this in mind, we looked at the global widefield activity to see if particular areas stand out for broadband spatial frequency representation. And although we did observe a slightly increased activity in V1, this was not as impactful as for the broadband orientation. Interestingly, when the stimulus was enriched in both orientation and spatial frequency information at the same time, it showed increased V1 activity, that was not as well visible in the 2-photon imaging.

Electrophysiological recordings also revealed that the mixed stimulus strengthens neuronal responses, especially in the deeper layers. This increase in response strength was strongest in the optical SC. The weak impact of spatial frequency bandwidth enrichment in the supragranular V1 layers and somatic spiking along with the stronger impact of this broadening in the widefield and LFP signals might indicate a connection to the dendritic signals of layer 5 neurons and overall population activity that represents additional dendritic and axonal signals. The impact of broadband spatial frequency stimuli on V1 could, therefore, be subtler than what single-cell activity measures used here can detect.

The increased preference in the SC for the mixed stimulus could suggest that less-predictable and more salient stimuli recruit SC circuits more strongly than the visual cortex. At the same time, the low impact of the spatial frequency rich stimulus could be in connection to the fact that the visual system might compensate for this richness by decorrelating the redundant information upstream to V1 supragranular layers. This reduction in representation can already be seen in the deeper V1 layers as well as the center stimulus responses (which represent feedforward information).

3.2.3 Broadband spatial frequency show no impact on visual perception.

The impact of spatial frequency bandwidth on visual perception has been tested with similar stimuli in humans (Simoncini et al., 2012). Broadband spatial frequencies showed to hinder perceptual speed discrimination in humans. This is potentially done via a gain normalization mechanism that reduces network responses to adjust for the large recruitment of the richly informative input signal.

Whereas, in the context of our discrimination task, the mice showed no change in their perception of orientations upon spatial frequency bandwidth enrichment. At the same time, they showed an improved perception of spatial frequencies upon orientation bandwidth enrichment due to a refinement in population code that allows for better separation of representations in the network. While this neuronal representation separability was not strengthened for spatial frequency information enrichment, neither did it get worse or hinder the discrimination of orientations in the behavioral task. The lack of information richness related neuronal recruitment could be due to species specific cortical organization and visual acuity ranges. Whereas the difference in perceptual impact could be task specific. For example: with broad spatial frequency content, more spatial scales carry the same orientation cue, thus more information about the orientations is being added. The mouse brain could pool across spatial scales to reinforce orientation signals at a population level to resolve the central orientation well enough. This applies to our orientation discrimination task specifically; speed discrimination could be affected by a different processing mechanism. Moreover, patterns of small moving bars drive the cortex more strongly than gratings, even at higher spatial frequencies (Dyballa et al., 2018). This points to a different mode of motion perception, which may help explain why we observed stronger responses to motion clouds with broad bandwidth (Dyballa et al., 2018; Muir et al., 2015).

3.3 General conclusion and outlook.

To summarize, we found that the mouse visual system chooses feature and context specific representation strategies to represent richly informative naturalistic visual input. The V1 network selectively adjusts its tuning based on the higher-order features of the visual input and base on the spatial arrangement of the orientation information in the scene.

When the orientation information is uniform across the space, hence highly predictable, the network sharpens the tuning by allowing the more sharply tuned neurons to dominate with their responses. It potentially achieves this through a horizontal and feedback driven activation of SOM interneurons that attenuate the responses of excitatory neurons. When the orientation information is diverse across the space, hence highly unpredictable, the network broadens the tuning by an increased and unselective recruitment of neurons that generates a stronger and more salient signal. It potentially achieves this through a horizontal and feedback driven activation of VIP interneurons that cause disinhibition by attenuating SOM interneuron activity. This salient signal enables better representations of highly unpredictable stimuli and thus an improved visual perception in mice. The engagement of the HVAs in a stimulus predictability dependent manner, places this mechanism in line with predictive coding (Rao & Ballard, 1999).

When the spatial frequency information enriches across space, it introduces a lot of redundant information about dense structures which are not very relevant for spatial orientation. The visual system decorrelates redundant information upstream to L2/3 of the mouse V1 to prioritize relevant and informative spatial frequencies and maintains a stable network activity by downregulating and upregulating differently tuned subpopulation of neurons, but keeping overall activity the same. Furthermore, since the predictability of the stimulus is not heavily impacted by this enrichment in spatial frequencies, surround suppression still plays a role in keeping the responses limited.

When both the orientation and spatial frequency information is diverse across the space, hence most unpredictable case, the saliency of the stimulus and higher order textures become more relevant than the decorrelation of the redundant information upstream V1 L2/3. Hence, the increased activation of optical SC, and deeper V1 layers. The lower impact of the orientation tuning regressor in the multiple regressor

LME for the mixed response modulation could imply an even broader and unspecific recruitment than the broadband orientation.

Placing all this within the big picture of nature visual processing: The relevant information content, salience and predictability of a natural scene defines the mode employed for representing it in the visual cortex. Highly unpredictable contexts drive a stronger signal for better representation, whereas highly redundant stimuli are decorrelated and encoded in a more efficient and sustainable way.

Though this work has revealed a few interesting mechanisms on natural image processing in the cortex, a lot of follow up experiments are required to give experimental validity to the mechanisms proposed here.

Awake recordings of behaviourally engaged mice are crucial to the follow up questions. Prolonged imaging of naïve, in-training and trained mice could reveal interesting representation changes that follow task-relevant stimulus contexts. Neurons are assigned a receptive field of a given size and location based on their interactions in the network, their excitatory and inhibitory inputs, be these consequential (driving a response) or modulatory (priming for a response, modifying the response amplitude etc.). Here, we show that at a network level, established properties like tuning sharpness and surround suppression regime are dynamic and context dependent. Since, it has previously been shown that individual neurons can change their neuronal tuning based in an experience dependent manner (Kreile et al., 2011), it would be interesting to see how much of this is stable and how much is plastic at a single neuron level.

Additionally, the central orientation used in this study was vertical (0°) which is one of the two best represented orientations in the V1 population tuning. It would, therefore, be important to see how much the broadband orientation enrichment effect is influenced by this over-representation. Thus, broadening from an oblique central orientation or testing a battery of central orientations (for which the tuning curve of the neurons has been characterized) could be instrumental in parsing the different aspects of internalized nature statistics representation. Even more provocatively one could ask how development affects this by comparing the orientation enrichment in mice from visually enriched or visually deprived rearing conditions.

While we focused on the immediate neural responses by showing brief stimuli, natural vision often involves sustained exposure to broadband information. Prolonged versus repeated broadband stimulation may engage different adaptation or facilitation mechanisms that affect the temporal response dynamics in different cortical layers and visual areas. For example, neurons in HVAs or in deeper layers or might show more facilitation or delayed suppression depending on stimulus predictability or task relevance.

Considering one would expect this suppression release to be negatively affected by anesthesia or optogenetic inhibition of HVAs, these two experiments would be crucial in giving experimental validity to the expectation that feedback signal influences surround modulation in V1 for the different information richness contexts. Additionally, laminar electrophysiology or L1 imaging of dendritic activity could reveal a lot here.

In line with this, it would be important to look at the axonal terminal from HVAs in or the neuronal responses in the HVAs, especially in lateral areas like LM which stood out in the WF imaging. One immediately relevant question to address there would be whether surround suppression plays out similarly in these areas.

The proposed neuronal circuitry also needs to be validated through simultaneous imaging and selective manipulation of SOM, PV and VIP and excitatory neurons with the help of Cre-dependent constructs. Additionally, specific inhibition of HVAs while imaging SOM could also help parse how much of the surround suppression comes from feedback and how much from local recurrent activity.

The preference of SC neurons for the highly unpredictable and salient mixed stimulus also opens the door for future explorations of deeper cortical and subcortical processing of broadband naturalistic stimuli.

One of the technical limitations mentioned in the discussion is that single cell level somatic responses we have measured are limited in exposing the subtle spatial frequency enrichment effect. A first approach to bypass this would be by focusing on the L5 apical dendrite activity in L1 to see if the widefield effects can be confirmed.

Furthermore, it might be of interest to test the impact of spatial frequency information richness on speed discrimination like in the human study (Simoncini et al., 2012). Broadband spatial frequencies enhanced visuomotor reflexes while hindering

perceptual speed discrimination in humans, while having no effect on the ability of a mouse to discriminate orientations. This would parse how much of the effect difference is context dependent and how much is species dependent. Translating our behavioral task in humans could also be of interest. Another interesting enrichment to the behavioral task would be implementation of the most unpredictable case, the mixed stimulus, to see how much the spatial frequency component hinders the orientation enrichment effect in mice or humans.

Broadband spatial frequency stimuli currently contain more of the less informative, lower spatial frequencies, due to their power law distribution. An interesting question would be what happens when the stimuli are tailored to broaden unilaterally and only introduce higher spatial frequencies. Would then a reduction in information decorrelation and increase cortical activation. This stimulus is would not be representative of nature, but an exception that proves the rule. A different engagement mode upon this non-natural stimulus enrichment, would add to the arguments that what has been observed here about spatial frequency enrichment is in accordance to the evolution of the mouse visual system for efficient representation of natural visual statistics.

In natural scenes, orientations are diverse, broadband and they tend to overlap and cluster around salient objects, giving rise to complex receptive fields in cortical neurons. Rarely do these neurons have to deal with isolated gratings of a single orientation and they often respond more strongly when features are embedded in natural contexts rather than in artificial contexts (Felsen et al., 2005). Understanding how the brain encodes such complex inputs is essential for uncovering the core principles of visual computation. Our findings show that broadband orientation motion cloud stimuli offer a powerful, yet parametrically safe, approach to explore these principles.

4. METHODS

***This section draws on material from Balla et al., Nature Communications* 16, 58003 (2025), licensed under CC BY 4.0.**

4.1 Experimental animals and preparation.

The model animal used for the experiments in this thesis was the mouse (*Mus Musculus*). All experiments performed on the mice were designed and executed in accordance with the German animal protection law and local ethics committee (LANUV, NRW) under the study protocols 84-08.04.2016.A357 and 81-02.04.2021.A021. The mice were selected to be between the ages of 14 and 30 weeks. They were kept on a reverse light cycle (light/dark cycle: 12/12 hours) with regulated temperature and humidity conditions (23 ± 2°C, 55 ± 5 % air humidity). For the 2-photon imaging experiments, six male and three female C57BL/6J (Charles River, Germany) mice (Fig. 7 - 10 show data from different days where these 9 mice were randomly chosen to run either the orientation or the spatial frequency protocol) and five 2Niell/J (B6; DBA-Tg(tetO-GCaMP6s)2Niell/J, Charles River, Germany) females (Fig. 12 and 14) were used. For the widefield imaging experiments, two male and two female 2Niell/J mice were used. For the electrophysiological recordings, three 2NiellJ males were used. As far as the behavioral discrimination tasks: three female and three male C57BL/6J mice were used for the spatial frequency discrimination task and six male C57BL/6J mice were used for the orientation discrimination task. The twelve mice used in the behavioral task were water-restricted during their respective experimental periods. They received water ad libitum on the weekend when no experiments were performed and received at least 1.5 ml of water on experimental days. They had access to food ad libitum and were weighed and checked for their health status before the start of each behavioral session.

4.1.1 Behavioral monitoring.

The mice were acclimatized with human interaction through daily handling and eventually tube habituation for a duration of one to three weeks prior to experiment start. Upon asserting that the animal was at ease in the tube, habituation to the 2-photon setup took place. During all imaging experiments, the mice were kept head-fixed and in the tube and their general behavior was observed using a webcam (Microsoft LifeCam Studio, Microsoft).

4.1.2 Surgical procedures.

For the cranial window implantation and viral injection, the mice were weighed and given a subcutaneous injection of Carprofen (4 mg/kg, Rimadyl, Zoetis GmbH) and Buprenorphine (0.1mg/kg, Buprenovet sine, Bayer Vital GMBH) for analgesia, 20 minutes before the surgical procedure. They were then anesthetized using (1%-2.5%) isoflurane in oxygen and placed in a stereotaxic frame with stabilized and monitored body temperature (37°C). Eye ointment (Bepanthen, Bayer Vital GmbH) was placed on the ipsilateral eye and ophthalmic gel in the contralateral to keep them moisturized but also allow intrinsic imaging to identify the location of V1(Andermann et al., 2011; Garrett et al., 2014; Kalatsky & Stryker, 2003; Roth et al., 2012). Following a local injection of Bupivacaine (0.08ml of 0.25% Bucain 7.5mg/ml, Puren Pharma GmbH), the skin was incised and pushed aside to reach the skull around the position of V1, based on anatomical coordinates (3.5 mm posterior and 2.5 mm lateral from bregma). The skull was then carefully thinned with a dental drill to gain optical access to the cortex for intrinsic imaging (see section 'Intra-surgical intrinsic imaging'). A 4-mm-wide circular craniotomy was then performed over the center of V1. In the C57BL/6J mice for 2-photon imaging, a viral vector (AAV9.Syn.GCaMP6f.WPRE.SV40; Addgene, viral titer ~10^{13} vg/ml, diluted 1:10 in phosphate-buffered saline (PBS)) was then injected through a thin glass pipette with a micropump (UMP-3, World Precision Instruments). In each mouse, we injected 250 µl at 100 nl/s in two separate V1 locations (~500µm apart from each other) to obtain an even expression GCaMP6f throughout the area (We waited around three weeks for stable expression before imaging). The dura was kept moist with PBS solution throughout the injection. After viral injections, a 4-mm-wide round coverslip was placed inside the craniotomy and fixed with light-curable dental cement (DE Flowable composite, Nordenta). For electrophysiology experiments, we used coverslips with a small opening (~0.2 mm) that was covered with silicon to allow access of the Neuropixels probes to the cortex. A metal head holder was then positioned and attached with dental cement to allow subsequent head fixation and imaging in the 2-photon microscope setup. In the 2Niell/J mice for the 2-photon experiments, the head-bar placement followed immediately after the window placement, and the location of V1 was identified later with widefield field sign mapping (see section 'Widefield imaging'). The custom stainless steel headbar piece was held in place with quick adhesive cement (Superbond C&B, Sun Medical CO., LTD) and fast-curing acrylic resin (Jet Denture repair, Lang Dental INC). For widefield imaging, 2Niell/J mice were similarly prepared, but a skull clearing procedure was performed instead of the cranial window (Z. V. Guo et al., 2014): the skull was thoroughly cleaned and covered with a thin, transparent layer of cyanoacrylate (Zap-A-Gap CA+, Pacer technology) to obtain optical access to the cortical surface (Couto et al., 2021).

After the surgery, all animals received the follow up injections of Buprenorphine and were kept over a heating blanket during their immediate recovery. They also received Buprenorphine (0.009 mg/ml, Buprenovet sine, Bayer Vital GmbH) and Enrofloxacin (0.0227 mg/ml, Baytril 5%, Bayer Vital GMBH) in their drinking water for three days. The recovery time was at least one week before starting animal handling. See Fig. 22 for a schematic description of all the steps described above.

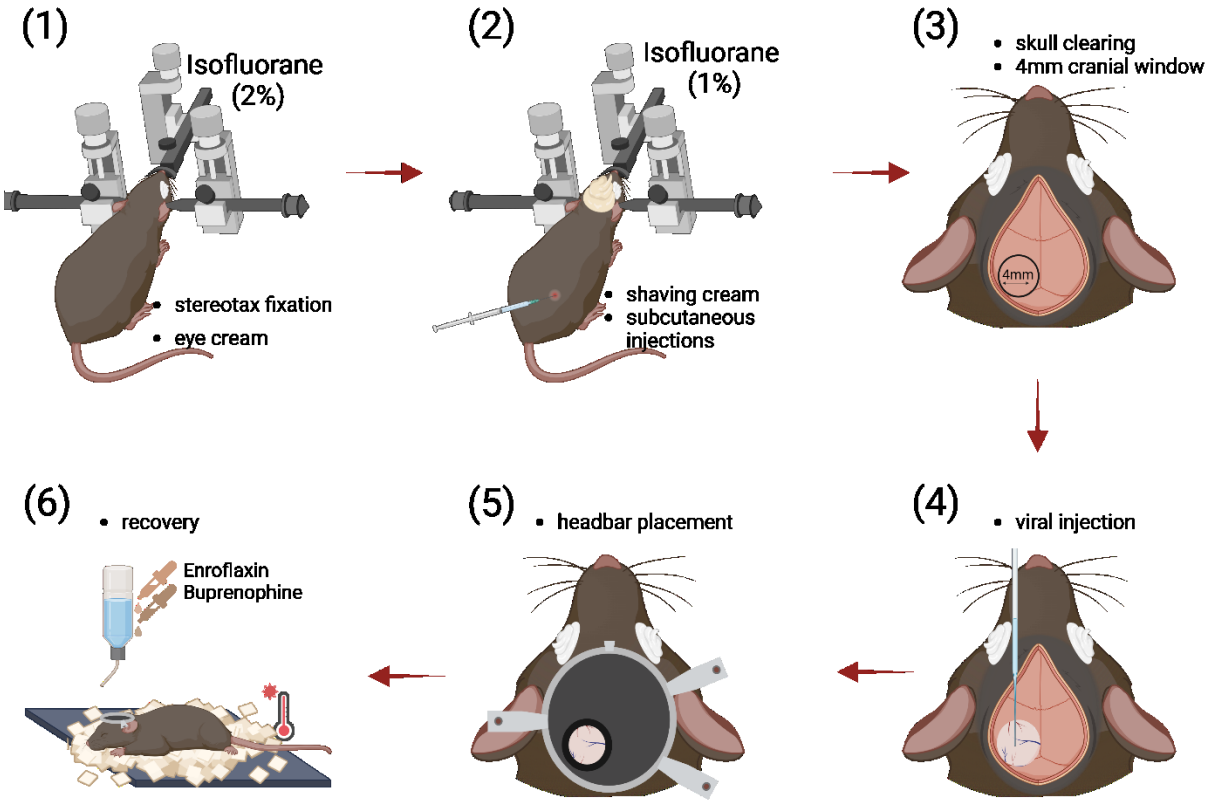


Figure 22. Schematic of the steps of the surgical preparation for imaging and headbar placement.

Schematic of the main surgery steps. Created in Biorender Balla, E. https://BioRender.com/g05l789 (2025).

4.1.3 Intrinsic imaging during surgery

During cranial window surgeries, retinotopic mapping was performed using intrinsic signal optical imaging to functionally confirm the location of V1. Here, we presented continuous periodically drifting bars with a high contrast flickering checker pattern (Garrett et al., 2014; Marshel et al., 2011). Stimuli were created using Psychtoolbox (Brainard, 1997; Pelli, 1997) in MATLAB (MATLAB R2019b, MathWorks) and composed of a static checker pattern with a

patch size of 20° and contrast reversal at 3Hz. This pattern was masked and only visible through a 15°-wide bar aperture, moving over the screen with a temporal period of 18 seconds in all four cardinal directions in randomized order. Stimuli covered roughly 120° horizontally and 80° vertically and were pre-rendered in advance. A spherical correction was applied to compensate for distortions due to the presentation on a flat monitor (Marshel et al., 2011). The monitor (BenQ XL2420T) was positioned at an angle of 20° to the midline and tilted 10° above the animal. The eye was positioned at the horizontal center of the monitor, 30 mm above the lower edge of the monitor, and the stimulus presentation was adjusted accordingly. A vessel image for alignment was captured under green-light illumination (peak ~525nm, green power LED, 15506, LUMITRONIX). Hemodynamic signals were acquired under red-light illumination (peak ~630nm, red power LED, 15607, LUMITRONIX). Images were captured using a tandem-lens macroscope composed of a 25 mm objective (DO-2595, Navitar) above the subject and a 50 mm objective in front of a CMOS camera (acA1920-155um, Basler). In between, a removable bandpass filter was used to isolate reflected red light (630 nm – 760 nm, G384082035, Qioptiq). Phase maps of horizontal and vertical retinotopy were generated online using custom software based on previously published analysis (Kalatsky & Stryker, 2003). Visual field sign maps, delineating the precise borders of V1 to higher visual areas, were then computed after imaging was completed (Garrett et al., 2014). During intrinsic imaging, anesthesia was kept light, just above 1% isoflurane in oxygen. Throughout retinotopic mapping, the state of anesthesia was closely monitored using an additional camera viewing the animal, as well as screening for global drifts in the hemodynamic signal.

4.2 Data collection approaches

4.2.1 2-photon imaging.

2-photon imaging was done in a custom-built setup with a resonant-scanning 2-photon microscope (CRS 8KHz, Cambridge Technology) and a femtosecond-pulsed TI: Sapphire laser (Mai Tai DeepSee, Spectra-Physics) tuned to 920 nm. Data acquisition and controls were performed with the MATLAB-based Scanimage package (Scanimage 2015) (Pologruto et al., 2003). Settings and laser power were kept constant across all mice and experiments. References from the vessel image, visual field sign maps, and/or injection site referencing were used to define the center of the primary visual cortex. Recordings were made from a single plane in layer 2/3 of V1 at a depth of 250-300 μm below the pial surface with a frame rate of 30 Hz. We used a 16x, 0.8 NA Nikon objective lens (CFI75 LWD 16X W) at a pixel

resolution of 512 x 512 pixels (575 µm × 575 µm). Motion Cloud stimuli from the orientation bandwidth expanding and the spatial frequency bandwidth expanding protocol (Fig. 7, 9, and 19, see 'Visual stimulation' section below) were presented at 17 cm distance from the right eye on a gamma-corrected LED-backlit LCD monitor (BenQ XL2420T, 24", 60Hz refresh rate, or Viewsonic VX3276-2K-MHD-2, 32", 60Hz refresh rate). The larger screen was used to increase the visual field coverage for center-surround stimuli, where we also used a sparse noise stimulus(Van Beest et al., 2021) to map the preferred RF locations of the individual neurons in the field of view of each recording. Two behavior cameras (Firefly DL Teledyne FLIR) recorded the animal from different angles throughout the sessions for movement and eye tracking.

4.2.2 Widefield imaging.

Widefield Ca^{2+}-imaging was performed using a custom tandem-lens macroscope(Ratzlaff & Grinvald, 1991), consisting of two 85 mm objectives (Walimex Pro 85mm f/1.4 IF; Walimex) using a sCMOS camera (Edge 4.2, PCO) through the cleared intact skull. Frames were acquired using the Python-based software package Labcams (Version 0.2 https://github.com/jcouto/labcams, by Joao Couto) at a framerate of 30 Hz with an effective resolution of 512 x 512 pixels with 4 x 4 spatial binning. Frames were acquired under alternating illumination using a blue LED (470 nm, M470L3, Thorlabs) and a violet LED (405 nm, M405L3, Thorlabs) with a 405 nm excitation filter (#65-133, Edmund optics). The excitation light was collimated using adjustable collimator lenses (SM2E, Thorlabs). Using a dichroic mirror (no. 87–063, Edmund optics), both excitation light paths were merged and then reflected onto the brain surface using a second dichroic mirror (495 nm long-pass, T495lpxr, Chroma). GCaMP6s fluorescence signals were isolated using a 525 nm emission filter (86-963, Edmund optics) in front of the camera. Images acquired under violet illumination captured calcium-independent fluorescence at the isosbestic point of GCaMP(Lerner et al., 2015). Therefore, by subtracting the linearly-rescaled calcium-independent signal from the calcium-dependent signal acquired under blue illumination, we could remove the intrinsic signal due to hemodynamic fluctuations. Widefield imaging data were motion corrected using a subpixel image registration routine(Guizar-Sicairos et al., 2008) was performed separately for frames acquired under blue or violet illumination. Then, a singular value decomposition (SVD) was used to reduce the dimensionality of the data and computational cost associated with the analysis(Musall et al., 2019). All subsequent analysis across time was performed on the 500 components that described the highest variance in the imaging data. Furthermore, a zero-phase, second-order Butterworth filter was used to remove slow signal drifts below 0.1 Hz

before performing the hemodynamic correction. All subsequent analyses were based on this low-dimensional and hemodynamic-corrected signal.

All recordings were aligned to the Allen CCF based on anatomical landmarks, and the location of V1 and higher visual areas (HVAs) was additionally confirmed using retinotopic mapping as described above(Marshel et al., 2011; Zhuang et al., 2017). Due to the faster neural signals with GCaMP-related widefield imaging, we used a faster movement of the bar aperture (2 s instead of 18 s per period with 20 repetitions for each direction). The same retinotopic mapping procedure was also used for mice in the 2-photon and electrophysiology experiments, where the imaging was done through the cranial window (Fig. S6a).

4.2.3 Electrophysiological recordings.

Electrophysiological recordings were done with Neuropixels 1.0 Probes(Jun et al., 2017) in head-fixed awake mice freely running on a wheel. For each mouse, we performed four recordings on subsequent days. To later recover the probe position in each recording, the probes were painted with DiD cell labeling solution (Invitrogen V22887) before each recording. To record neural responses to motion cloud stimuli across all cortical layers, we targeted the center of V1 (4 mm posterior and 2.5 mm lateral from bregma) and inserted the probe perpendicular to the cortical surface (6 recordings in 3 mice) to a depth of 2 mm. To simultaneously record from V1 and SC, the probes were angled to 35° elevation and inserted to a depth of 3.6 mm. Neuropixels were positioned with a high-precision micromanipulator (uMp-4, Sensapex) at a speed of 10 µm/s. In each recording, the insertion point was moved by ~100 um in either mediolaterally or rostral-caudal direction from the target coordinate to ensure that fluorescent electrode tracts could be later related to the probe position in individual recording sessions (Fig. S5a and S6b). To determine the surface of the SC, we took advantage of the visible gap in visual responses between the cortex and SC and determined the SC surface at the depth at which we could detect visual LFPs after this gap. In agreement with our histological results, this was usually at a depth of 1.63±0.16 mm.

All visual stimuli were presented at a 17 cm distance from the right eye on a gamma-corrected LED-backlit LCD monitor (Viewsonic VX3276-2K-MHD-2, 32”, 60Hz refresh rate). Similar to the 2-photon experiments, we then presented visual stimuli from the orientation band expanding and spatial frequency band expanding protocols (see ‘Visual stimulation’ section below). Data acquisition from the Neuropixels probe was done using SpikeGLX v3.6 (https://github.com/billkarsh/SpikeGLX, Bill Karsh). High-pass filtered signals were recorded from 384 channels at 30 kHz, and low-pass filtered signals were recorded at 2.5 kHz. The data was then time-shifted and median-subtracted with CatGT 3.9. Channels that were broken or outside the brain were detected using SpikeInterface(Buccino et al., 2020) and removed from

further analysis. The data was then spike-sorted with Kilosort 2.5 and manually curated using phy 2.0b1 (https://github.com/kwikteam/phy). The criteria for manual curation were consistent spike waveforms across at least three recording channels, stable spike amplitudes across the recording, and a visible gap in the center of the auto-correlogram. In addition, we imposed a cutoff for neurons with inter-spike-violations above 0.5, using the contamination rate method(Hill et al., 2011). Sorted data were analyzed using custom MATLAB code.

After all electrophysiological recordings were complete, the mice were first perfused with cooled PBS and subsequently with 10% formalin. After storing the brains in sucrose solution with increasing concentration from 10-30%, 50 µm-wide brain slices were cut using a Cryostat (CM3050 S, Leica). Images of the brain slices were made using a fluorescent microscope (BZ x810, Keyence, Germany).

4.3 Visual stimulation.

Motion cloud stimuli were generated using the Motion Cloud Package(Leon et al., 2012) and adapted for display via custom MATLAB code (MATLAB R2019b, MathWorks). The package follows a linear generative model for creating a dense mixing of localized moving gratings with random positions based on predefined parameters. To ensure that neural responses were not strongly driven by the individual properties of these random phase motion textures, 10 renderings (300 frames each) of each stimulus were generated for each parameter combination. 20 repetitions of each rendering were then presented in a pseudorandom sequence movie where stimulus presentation order was shuffled with a restriction that prevents consecutive presentation of the same stimulus category. Motion cloud stimuli were either enriched in orientation bandwidth (narrow: 5°, mid: 22.5°, broad: 45°; around a central orientation of 0°; see also Fig. 7) or in spatial frequency bandwidth (narrow: 0.004 cpd, mid: 0.04 cpd, broad: 0.4 cpd; around a central spatial frequency of 0.04 cpd; see also Fig. 9). To test the discrimination of different central orientations or spatial frequencies (Fig. 19, 21) the central orientation was either 0° or 90° and the central spatial frequency either 0.04 cpd or 0.16 cpd. All motion clouds were presented on the screen, drifting from nasal to temporal (or orthogonal to the central orientation) with a temporal frequency of 1 Hz. For the orientation and spatial frequency bandwidth protocols, the stimulus duration was 5 seconds, followed by a 5-second inter-stimulus interval during which a mean gray screen was shown. An overview of all parameter combinations and protocols is shown in Table 1.

	Central Orientation (°)	Orientation bandwidth (°)	Central Spatial Frequency (cpd)	Spatial Frequency bandwidth (cpd)
Orientation band expanding	0	5, 25, 45	0.04, 0.16	0.004
Spatial frequency band expanding	0, 90	5	0.04	0.004, 0.04, 0.4

Table 1. Specific Parameters for the orientation and spatial frequency band expanding protocols.

To investigate the effect of the orientation bandwidth band expanding (orientation band enrichment protocol, top row), the central orientation (first column) and spatial frequency bandwidth (last column) were kept constant (0° and 0.04cpd, respectively), while different combinations of orientation bandwidth (second column: 5°, 25°, 45°) and central spatial frequency (third column: 0.04 cpd and 0.16 cpd) were presented. This resulted in a total of 6 different stimulus parameter combinations. The same principle was used for the spatial frequency band expanding protocol (bottom row).

To combine changes in both orientation and spatial frequency bandwidth (Fig. 14, 18; 'mixed' condition), we used motion cloud stimuli with a central spatial frequency of 0.04 cpd, a central orientation of 0° and a temporal frequency of 1 Hz, and then increased the bandwidth of either spatial frequency, orientation or both. The 'narrow' case consisted of 0.004 cpd spatial frequency bandwidth and 5° orientation bandwidth and served as the baseline for the narrow bandwidth of both features. In the 'SF' case, the spatial frequency was increased to 0.4 cpd while the orientation bandwidth remained at 5°. In the 'ORI' case, the spatial frequency bandwidth remained at 0.004 cpd, and the orientation bandwidth was increased to 45°. Lastly, in the 'mixed' condition, both spatial frequency and orientation bandwidth were increased to 0.4 cpd and 45°, respectively. For each stimulus condition, we generated eight separate stimulus renderings to reduce the risk of over-emphasizing neural responses to complex features that could occur in a specific rendering. Each rendering had a total of 7 presentations as full-field stimulus (56 presentations for each full-field stimulus).

To test for the impact of center surround suppression, we first used sparse noise stimulation to identify the receptive field locations of the neurons in each recording session. The stimulus in each frame was composed of four non-adjacent 12°x12° white squares over a black background(Van Beest et al., 2021). The grouping of the squares was pseudo-randomized to ensure that each location would be presented 20 times in total. Each presentation lasted for

250 ms with an inter-stimulus interval of 750 ms. For each field of view, we then performed an imaging session with sparse noise stimuli to identify the receptive fields (RF) of all imaged neurons (see section 'Data analysis and statistics'). After the imaging session, we then created a new stimulus set where all motion clouds were covered by a gray mask, and a circular, 15° transparent aperture was placed in the center of the mean RF of the imaged population (Fig. 14b, 'Center RF,' see also Fig. 15). In a second recording session, we then returned to the same field of view (based on the vessel pattern and visually comparing the location of the imaged cortical neurons) and repeated the sparse noise stimulation. Subsequently, we presented the full-field and center RF stimuli for the narrow, SF, ORI, and mixed conditions, as described above. For both the full-field and center RF conditions, we presented 56 stimuli, respectively, in pseudo-randomized order. The stimulus duration was 2 s, followed by a 4 s gray screen. To test for orientation tuning (Fig. 14), we showed full-field drifting square wave gratings at five different orientations (-45°, 22.5°,0°, 25°, 45°; 0.04 cpd spatial frequency; 1 Hz temporal frequency) in pseudo-randomized order. The stimulus duration was 1 s, followed by a 2 s inter-stimulus interval where a gray screen was shown, and each stimulus was repeated 20 times.

4.4 Touchscreen chamber experiments.

Behavioral experiments were performed in a custom-built operant conditioning touchscreen chamber, using a 10" screen for visual stimulation (Elo Touch 1002L monitor, running at 60Hz)(Smith & Edgar, 1990). All animals performed the experiments under a water restriction regime and received at least 1.5 ml of water per day to maintain their general well-being. They received water *ad libitum* on the weekends and had access to food *ad libitum* throughout the experiment. All mice were weighed daily and monitored before and during the behavioral experiments. In the first phases of pre-training, the animals were habituated to the touch screen chamber and learned to associate a green light with the availability of a reward. After the animals were able to collect at least one reward per minute, they proceeded to the next phase, where they had to first touch the screen to receive a reward, which was still indicated by a green light. Again, after the animals performed the sequence of touching the screen and collecting a reward at least once a minute, they were moved to a more specific training regime to perform either the orientation or spatial frequency discrimination task. All trained mice passed the initial criteria and could be trained on the discrimination tasks.

4.4.1 The orientation discrimination task.

The *orientation discrimination task* started with the training for easy discrimination, where the animals had to discriminate a horizontal (target) from a vertical (distractor) sine-wave grating with a spatial frequency of 0.04 cpd. Each stimulus was shown through a circular aperture with a diameter of 7.5 cm, and stimuli were presented left and right from the screen center, with a distance of 10 cm between the stimulus and the screen center. After the animals reached a performance of at least 80% correct responses, we proceeded with a staircase procedure. In this staircase procedure, the difficulty level was dynamic. Every correct response was followed by a decrease in orientation difference of 3°, while an incorrect response was followed by an increase in orientation difference of 8°, leading to a target performance of 72.7%. After the animals could perform the staircase procedure, we obtained one orientation discrimination threshold per session, revealing the minimal orientation difference between the target and distractor stimulus that was required for robust task performance. In randomized sessions, motion clouds with different spatial frequency bandwidths were used. Only one spatial frequency bandwidth was used per session to test how the discrimination threshold would be affected by a given spatial frequency bandwidth. To later compare the perceptual impact of different spatial frequency bandwidths, all trials with an orientation difference between 40° and 50° were combined and averaged over all sessions.

4.4.2 The spatial frequency discrimination task.

The *spatial frequency discrimination task* followed a similar training regime. After the animals performed sufficiently in the pretraining (Fig. S7), they were trained to discriminate between two (0° central orientation) motion clouds of different spatial frequencies. Throughout pre-training, the animals had to discriminate different spatial frequency pairs (0.04 cpd vs 0.06, 0.12 or 0.16 cpd; Fig. S7). The animals showed a stable discrimination between 0.04 and 0.16 cpd, after which different orientation bandwidths were introduced. The different orientation bandwidths were used within the same session because no adaptive staircase procedure was used in these experiments. To ensure that animals did not develop an intrinsic side bias during the discrimination tasks, we used a bias correction routine(Wiesbrock et al., 2022). In the initial ten trials, the position of the correct stimulus was randomized, and the bias correction was subsequently used to correct for direction and repetition bias. In case a bias was detected, the stimulus position was chosen to increase the probability of showing the target stimulus on the non-preferred side. If a repetition bias was detected, the probability of changing the target stimulus position from one trial to the next was increased. When no bias was detected, the position of the target and distractor stimulus was chosen randomly.

4.5 Data analysis and statistics.

All data analysis was done in MATLAB.

4.5.1 Widefield.

When computing V1 responses in the widefield, we first determined the 50% region of V1 that responded the strongest across all presented stimuli per mouse. This was needed to prevent obtaining artificially lower response magnitudes in V1 due to integrating too many pixels from medial V1 that showed much weaker visual responses. For Fig. 18c, V1 activity was averaged over this subregion of V1 and individual trial responses were pooled across mice.

4.5.2 Electrophysiology.

To quantify the electrophysiological spiking responses of individual neurons to broadband stimulation (Fig. 18f-h) we computed the area under the receiver-operator characteristic curve (AUC) between the baseline and the stimulus period. This allows us to obtain a measure of how reliably a given neuron responded to each stimulus, regardless of its overall firing rates. As with the 2-photon data, we first computed the mean firing rate during the stimulus and a 1-second baseline period for each stimulus presentation. To identify stimulus-responsive neurons, we then used a one-sided Mann-Whitney-U test to assess if the mean firing rate across all stimulus presentations was significantly larger than zero (p-value <0.05). For all stimulus-responsive neurons, we then computed the AUC between the baseline and stimulus periods to quantify how reliably they responded to each stimulus condition. An AUC of 1 indicates a perfect separation of the activity in the baseline and the stimulus period and an AUC of 0.5 suggests complete overlap.

4.5.3 2-photon imaging.

After each recording, 2-photon imaging data were motion corrected and individual neurons were isolated, using the Suite2P package with model-based background subtraction[61]. Suite2P was used to perform rigid motion correction on the image stack, identify neurons, extract their fluorescence and correct for neuropil contamination. ΔF/F traces for each neuron were then produced using the method of Jia et al.[62], skipping the final filtering step. To compute the stimulus response amplitude (ΔF/F) of individual neurons in the 2-photon data we computed the median across all trials for the difference between the fluorescence during the first 2-3 seconds of the stimulus and a preceding 1-s long baseline window for each stimulus presentation.

To identify neurons that were responsive for a specific visual stimulus, we used a one-sided Mann-Whitney-U test to assess if ΔF/F across all stimulus presentations was significantly larger than baseline (p-value < 0.05).

To identify the receptive field of neurons from the sparse noise paradigm, we also computed the ΔF/F for each neuron to all sparse noise stimuli that contained a white square in a specific part of the screen. This resulted in a set of neural responses for each 12° x 12° square on the screen and we used a one-sided Mann-Whitney-U test to identify significant responses to visual stimulation on a specific part of the screen. For significantly responding neurons, we then inferred the center of the receptive field as the stimulus location that resulted in the strongest neural responses.

We used DeepLabCut(Mathis et al., 2018) V2.3.9 to analyze the pupil position (see Fig. 15). Eight reference points around the lids and one pint at the center of the pupil were labeled and tracked. We binned the pupil movement range area (4 pixels/bin) and then calculated the mean frequency at which the center was located within a bin during the stimulus period.

To compute the orientation tuning for different neurons, we computed the ΔF/F in response to different grating orientations and used a one-sided Mann-Whitney-U test to identify neurons that significantly responded to any grating stimulus. Beside this, we did a linear interpolation of the tuning curves and inferred the preferred grating orientation of responsive neurons as the orientation that induced the strongest neural response. Only neurons with a clear orientation tuning, i.e. a difference between peaks and troughs above the median of the peak and trough distribution of the whole population, were selected for the multiple regressor model (Fig. 14h, i).

4.5.4 Statistics.

Throughout the results section, we have reported mean ± s.e.m. When reporting the normalized responding cell count for each stimulus per session, we normalized the cell count per session by the number of neurons that responded to the lowest respective bandwidth. We then pooled the cell counts from all recordings to test for significant differences in the number of responsive cells for different feature bandwidths. Since recordings from the same animals are not fully independent variables, we used a linear mixed-effects (LME) model (using the fitlme function in MATLAB) to determine significant differences between groups wherever possible (normally distributed data). Here, we included the stimulus and animal identity for the fixed and random effect coefficients, respectively. This allowed us to test if the stimulus identity significantly explains differences between two stimulus conditions while controlling for potential differences between animals. To test for significance, we performed a likelihood ratio

test between a mixed-effects model that only contained animal identity versus a full mixed-effect model that contained both stimulus and animal identity. The reported p values and T statistics, therefore, indicate whether the addition of the stimulus regressor added significant information beyond the animal identity to explain differences between the tested conditions. We used this approach to test for significant differences whenever samples (individual neurons or recordings) were pooled across multiple animals (Fig. 7e; Fig. 9e; Fig. 12b-e; Fig. 18c-h; Fig. 19b-d, f-h). Additionally, for testing the effect of orientation tuning versus surround modulation in Fig 14 g-i, we upgraded the model to a multiple-regressor linear mixed effect model, whereby the orientation tuning, the center response, and the surround modulation effect were used as regressors to determine the impact of each regressor on the resulting changes.

The LME model formula was:

$$\mathrm{Rm}\,ij \;= \beta 0 \;+\; \beta 1\,\mathrm{Sm}ij \;+\; \beta 2\,\mathrm{Ot}ij \;+\; \beta 3\,\mathrm{Cr}ij \;+\; \mathrm{u}j \;+\; \epsilon ij \tag{1}$$

where Rm_{ij} is the response modulation for the *i*th observation within the *j*th animal, β_0 is the fixed intercept (the total mean response modulation across all animals and conditions), β_1 β_2 and β_3 are the fixed effect coefficients for the predictors: Sm (Surround modulation), Ot (Orientation tuning) and Cr (Center response). u_j is the random intercept for the *j*th animal, accounting for the variation between different animals, and ϵ_{ij} is the residual error term for the *i*th observation within the *j*th animal, which accounts for the variability not explained by the fixed effects or the random intercepts.

In cases where using the LME model for statistics was not possible due to the data being not normally distributed, such as for ΔF/F amplitudes of pooled cells (Fig. 7f-h; Fig. 9f-h; Fig. 14d-f; Fig. 21c-f), we used a Wilcoxon signed rank test and additionally did a mouse-wise comparison in Fig. 8 and Fig. 10.

If multiple comparisons were performed, for example, when comparing the number of responsive cells in the mid and broad-bandwidth to the narrow-band condition (Fig. 7e), we performed a Bonferroni-correction for the number of performed tests to adjust for a multiple-comparison bias. Here, we adjusted the α level to reach significance to α = 0.05 / n, where n was the number of performed tests. In the above example, two tests were performed, so a significant difference required a p-value below α = 0.05 / n = 0.025.

To further test if individual neurons selectively responded to the narrow-, mid-, or broad-range of either the orientation or spatial frequency bandwidth (Fig. 8d, and 10 h), we computed a bandwidth selectivity index BW_SI that was computed as the response difference to a given

bandwidth (e.g., narrow-range, *Rn*) and the mean of the other two (e.g., mid- and broad-range, *Rm* and *Rb*), normalized by their sum.

$$BW_SI = \frac{Rn - \text{mean}(Rm, Rb)}{Rn + \text{mean}(Rm, Rb)} \quad (2)$$

The BW_SI is bounded between -1 and 1, with values > 0 indicating selectivity for a given bandwidth, such as the narrow range, while negative values indicate selectively weaker responses to that bandwidth. To test if the BW_SI for a given neuron could not be observed by chance, we performed a shuffle control. Here, we shuffled the labels for three bandwidth conditions and then re-computed the BW_SI to obtain a measure of possible BW_SI values that can occur due to random response fluctuations. For each neuron, we computed 1000 shuffled BW_SIs to obtain a random distribution and then tested if the real BW-SI was above the 95th percentile of the shuffle distribution. Neurons for which this was the case were counted as selectively responding to narrow, mid, or broad-range stimuli (outer paddle of the Euler diagrams in Fig. 8d, and 10 h). Neurons with negative BW_SI below the 5th percentile were counted as being mixed selective for the two bandwidths aside from the non-preferred bandwidth. For example, a neuron with a negative BW_SI below the shuffle control for the narrow-range bandwidth was counted as being selective for both the mid- and broad-range bandwidths (overlapping paddles of the Euler diagrams in Fig. 8d, and 10 h). Lastly, neurons without any selectivity beyond the shuffle control were counted as non-bandwidth-selective (the center region of the Euler diagrams in Fig. 8d, and 10 h).

Using the same approach, we computed an orientation modulation index OMI to quantify the difference in neural responses to broad versus narrow feature bandwidth (Fig. 19c, g).

$$OMI = \frac{R_{broad} - R_{narrow}}{R_{broad} + R_{narrow}} \quad (3)$$

where R_{broad} and R_{narrow} are mean responses to broad- and narrowband stimuli, respectively. The OMI is bounded between -1 and 1, with positive values indicating increased responses to broadband stimulation and negative values indicating reduced responses to broadband stimulation.

To determine the ability of each neuron to discriminate between the two different central spatial frequencies or between two central orientations, we also computed the AUC between these stimulus conditions (Fig. 19 c, g and Fig. 23). We computed the median response amplitude to all stimulus presentations for a given central feature (0.04 versus 0.16 cpd for spatial frequency and 0° versus 90° for orientation) and computed the AUC as a normalized measure

of response discriminability. Here, we used the absolute AUC as a measure of discriminability, regardless of the response preference for a given stimulus feature for a given neuron. This was done by computing the absolute AUC as

$$AUC_{abs} = |AUC - 0.5| * 2 \quad (4)$$

AUC_{abs} is bound between 0 and 1, with positive values indicating increased discriminability of neural responses between the two stimulus conditions. This is referred to as spatial frequency or orientation discrimination index in the Results section.

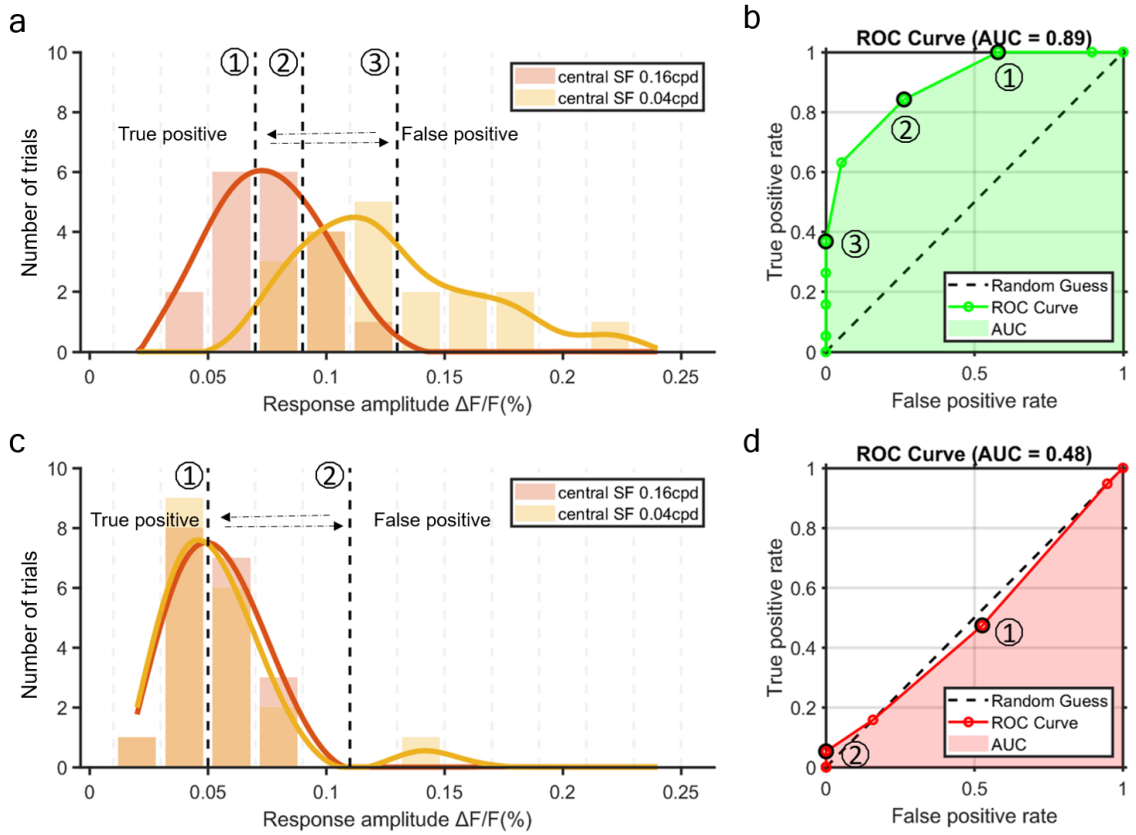


Figure 23. Example AUC for a strong and a weak central SF discriminator neuron.

Example central SF response distributions and corresponding ROC curves for a strong central SF discriminator neuron (top row) and a weak central SF discriminator neuron (bottom row). (**a**) Histogram of response amplitude distributions for the broad orientation stimulus condition (45° orientation bandwidth) with a low central SF (0.04 cpd, light orange, n = 19 trials) versus a high central SF (0.16 cpd, dark orange, n = 19 trials) for an example strong central SF discriminator neuron. Light and dark orange distribution overlap minimally (Wilcoxon ranked-sum test, p = 0.1 x 10^{-4}). (**b**) Receiver-operator characteristic (ROC) curve showing true vs. false positive rates for distributions in panel a. When response distributions are distinct at a given threshold, the true positive rate is high, and the false positive rate is low, moving the ROC curve away from chance (dotted line). Varying the threshold shifts ROC values, with AUC quantifying separability (0.5 = no discrimination, 0/1 = perfect discrimination). For responses to central spatial frequencies and to central orientations (Fig. 19), we used the absolute AUC (AUCabs = |AUC-0.5|*2) as a measure of discriminability. AUC_{abs} is bound between 0 and 1, with larger values indicating increased discriminability of neural responses between the two stimulus conditions. Black circles (1-3) on the ROC curve correspond to thresholds in panel a. The more rightward the threshold is placed, the more the corresponding ROC marker drifts to the left, decreasing the false positive rate but also the true positive rate at some point. In position 1, the true positive rate is 1 since all low central SF responses are right of the threshold. However, 58% of

high central SF responses are also right, giving a false positive rate of 0.58. Moving the threshold to position 3 reduces false positives to 0 but also lowers the true positive rate. An optimal separating point would be the threshold in position 2, where the best ratio between true positive (0.84) and false positive (0.26) rates is found. (**c**) Same a) for a non-discriminating neuron (see the large overlap between the distributions, Wilcoxon ranked-sum test, p = 0.81). (**d**) Same as b), for the non-discriminating neuron. Here, the orange distributions largely overlap, and true and false positive rates are symmetric, regardless of where the threshold is placed. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

4.6 Visualization.

All data figures were generated in MATLAB (MATLAB R2022a, MathWorks). Raincloud plots were generated with the Raincloud tool(Allen et al., 2021), and figures (Figures 1c, 2c, 3a, 4a, d, 5a, 7a, S4c, d, g, h, S9a) were created in BioRender(*Created in BioRender. Balla, E.*, 2025). For better visualization of the raincloud plots, outlier elements above or below3 median absolute deviations were excluded from the plots using the *rmoutliers* function in MATLAB. All values were included in the statistical analyses. Stars mark p-value ranges, whereby one to three stars respectively signify p-values smaller than 0.05, 0.01, and 0.001. If multiple tests were performed, significance thresholds were corrected for the number of comparisons (e.g., 0.05/n for one star).

4.7 Statistical analysis of natural images and motion cloud stimuli.

4.7.1 Orientation distribution analysis.

The ImageJ(Schneider et al., 2012) Directionality(Liu, 1991) plugin was used on a library(Tkačik et al., 2011) of nature photos (Fig. 7a, n = 1077 images) or a complete rendering of the motion cloud stimuli (Fig. 7b, n = 300 frames) to derive the distribution of directional changes in color intensity within a local gradient (5x5 Sobel filter is applied). Each image was loaded individually, and 180 bins of 1° each (range 1° - 180°) were selected. For each image, a data table of measured orientation contents in the given bin of orientations was saved. Finally, we calculated the mean and s.e.m. across all the individual images and plotted the respective curves (a light Gaussian fit, $\sigma = 2$ was applied) via custom MATLAB (MATLAB R2022a, MathWorks) scripts.

4.7.2 Spatial frequency spectrum analysis.

We analyzed the distribution of spatial frequencies for a set of images from either a library of nature photos(Tkačik et al., 2011) (Fig. 9a, n = 1077 images) or a complete rendering of the motion cloud stimuli (Fig. 9b, n = 300 frames). Here, we computed the 2D FFT spectrum for

each image individually. To describe the spatial frequency content of images independent of their orientation, we shifted the zero-frequency component to the center and averaged energies in the frequency domain depending on their Euclidian distance to the zero frequency. Units of the spatial frequency spectra were converted from cycles per image to cpd based on their presentation or acquisition. For motion cloud stimuli, we computed their angular size based on their presentation:

$$angular\ size\ =\ 2\ x\ arctan\left(\frac{monitor\ width}{2\ x\ distance\ to\ the\ eye}\right) \quad (5)$$

For the natural images, the angular size was derived based on the individual camera field angle of view of each album:

$$FOV\ =\ 2\ x\ arctan\left(\frac{sensor\ dimension}{2\ x\ crop\ factor\ x\ focal\ length}\right) \quad (6)$$

Finally, we calculated the mean and s.e.m. across all individual images and plotted the respective spectra.

4.7.3 Stimulus predictability analysis.

Spatial autocorrelations were computed for individual stimulus frames using the analysis provided by (Portilla and Simoncelli, 2000)(Portilla & Simoncelli, 2000). For the use of the Portilla-Simoncelli (PS) analysis, stimulus frames were cropped to a size of 256 by 512 pixels and spatially down-sampled by a factor of 2. Raw coefficient correlations presented in Fig. 14c were computed as the central sample of the real autocorrelation for a spatial neighborhood of 51 pixels. Maps of spatial autocorrelations were averaged over all frames in the stimulus library of a given condition and presented with the corresponding colormaps.

4.7.4 Stimulus separability.

To estimate how variable different realizations of the same stimulus condition are, we performed a principal component analysis across all texture parameters generated by the PS analysis (including pixelStats, pixelLPStats, autoCorrReal, autoCorrMag, magMeans, cousinMagCorr, parentMagCorr, cousinRealCorr, parentRealCorr, varianceHPR). We computed the principal components for individual frames and then averaged coefficients across frames within individual stimulus realizations. This resulted in average coefficients for

each of the eight realizations per stimulus condition, of which we present a projection onto the first two principal components (Fig. S3a).

4.7.5 Coefficient magnitude statistics.

We investigated if expanding orientation- and SF bandwidths result in the emergence of higher-order structures in the stimuli by computing the cross-orientation- and cross-scale energies(J. Freeman et al., 2013; Portilla & Simoncelli, 2000) (Fig. S3b, c). For this purpose, we computed the PS analysis using four orientations over six spatial scales for the steerable image pyramid decomposition. Cross-orientation energies are based on the “cousinMagCorr” parameter of the PS analysis, representing the correlations between sub-band magnitudes of different orientations within the same spatial scale(Portilla & Simoncelli, 2000). Here, we computed the average cross-orientation energy on the coarsest spatial scale between the 0° and the ±45° sub-band magnitudes, as these represent the dominant features in our stimuli. Cross-scale energies, given by the “parentMagCorr” parameter of the PS analysis, represent correlations of sub-band magnitudes with those of all orientations at the next coarser spatial scale(Portilla & Simoncelli, 2000). Here, we focused on the cross-scale energy of the 0° magnitude sub-band of the steerable image pyramid between the two coarsest spatial scales. For both cross-energy measures, we tested for statistical differences between the stimulus conditions, using pair-wise Wilcoxon–Mann–Whitney U tests with subsequent Bonferroni correction, comparing energies across stimulus realizations. For this, we averaged cross energies over frames belonging to the same stimulus realization. For visualization purposes, we displayed the cross energies of individual stimulus frames after removing outliers.

4.8 Model for neural responses of V1 neurons.

To explore if the increased neural responses and recruitment for larger orientation bandwidth stimuli can be explained by different orientation tuning of V1 neurons, we created a simple bank of Gabor filters model. To predict neuronal responses, we compute the response magnitude of individual Gabor filters using the imgaborfilt-function in MATLAB (MATLAB R2022a, MathWorks). The response magnitude of complex Gabor filter kernels represents the response of an ideal-phase filter; the model, therefore, assumes a uniform distribution of phase representation in V1. We also designed the bank of Gabors to match the known tuning properties of mouse V1. For this, all Gabor filters were generated with a spatial frequency

bandwidth of 2.38 octaves, matching previous reports(Niell & Stryker, 2008; Van Den Bergh et al., 2010). Further, we adjusted the spatial aspect ratio of the Gabor filters, determining the envelope ellipticity of the filter kernels, thereby affecting the orientation tuning width of a given filter. We estimated this parameter for Gabor filters with a spatial frequency of 0.04 cpd, presented with a corresponding full contrast, 0.04 cpd sine wave grating. To match the tuning width of ~22° half-width half-maximum previously reported for mouse V1 neurons(Niell & Stryker, 2008; Van Den Bergh et al., 2010), we used Gabor filters with a spatial aspect ratio of 0.55 (Fig. S1b). This represents an elongation of the supporting envelope in the direction of the parallel stripes of the filter. Furthermore, to capture a broad range of responses, we used Gabor filters tuned to 5 different spatial frequencies (0.01 cpd, 0.02 cpd, 0.04 cpd, 0.06 cpd, 0.08 cpd; Fig. S1c) within the predominantly represented range of spatial frequencies of mouse V1[2]. For each spatial frequency, we used filter kernels with 180 different orientations in 1° increments, resulting in a bank of 900 individual complex Gabor filter kernels that covered the full range of orientations. To compare and combine responses across Gabor filters with different spatial frequencies, we normalized Gabor filter responses to their optimal stimulus (a full contrast sine-wave grating of matching spatial frequency and orientation).

We then simulated V1 population responses to our motion cloud stimuli. To reduce the computational cost, we spatially downsampled stimulus frames by a factor of 4 and computed Gabor filter responses for each individual "location" within the central 30° by 30° patch (Fig. S2) around the stimulus center (37x37 sites in total) in the following denoted as such:

$$m(o, f, b, x, y) \mid o \in O, f \in F, b \in B, x \in X, y \in Y$$

$$O = (\{-45\,°, -22.5°, 0°, 22.5°, 45°\}; n_O = 5),$$

$$F = (\{0.01\ cpd, 0.02\ cpd, 0.04\ cpd, 0.06\ cpd, 0.08\ cpd\}; n_F = 5)$$

$$B = (\{narrow\ orientation\ bandwidth, broad\ orientation\ bandwidth\}; n_B = 2)$$

$$X = Y = \{1,2, \ldots ,37\}$$

(7)

Here, m represents the response magnitude of a Gabor filter with orientation tuning o and spatial frequency tuning f in response to a stimulus with orientation bandwidth b at the location (x, y) within the center of the stimulus frame. x and y denote indices of the 37x37 grid of locations corresponding to the 30° x 30° patch around the stimulus center. Subsequent model fits were generated based on the response at the stimulus center, either on its own or modulated by the average surrounding filter responses (Fig. 12, Fig. S1, and Fig. S2). The

Gabor responses at the stimulus center were obtained using the center position and averaging over the 5 Gabors with different spatial frequency tuning:

$$c(o,b) = \frac{1}{n_F} \sum_{f \in F}^{n_F} m(o,f,b,x_c,y_c)$$

(8)

To translate these normalized Gabor filter response magnitudes into ΔF/F values that matched the measured neuronal response amplitudes, we computed a scaling factor w_t (Fig. S2). To determine this factor, we first measured the median responses of V1 neurons with different orientation tuning to narrow and ORI stimuli (-45 °, -22.5°, 0°, 22.5°, 45°; Figure 12b, c) and then calculated the mean across all bins to equally weigh responses from differently tuned neurons. The average neural response amplitude was 1.622% ΔF/F. Second, we selected response magnitudes from Gabor filters with matching orientation tuning contained in set O and averaged responses across filters with these five orientation tunings, five spatial frequencies, and across stimulus conditions. This yielded an average normalized Gabor response magnitude of 0.1627 arbitrary units. From these two measures, we then derived the scaling factor:

$$w_t = \frac{1}{n_O n_B} \times \frac{\sum_{b \in B}^{n_B} \left(\sum_{o \in O}^{n_O} a(o,b) \right)}{\sum_{b \in B}^{n_B} \left(\sum_{o \in O}^{n_O} c(o,b) \right)} = \frac{0.01622\ \Delta F/F}{0.1627\ AU.} = 0.0997$$

(9)

Here, w_t represents the scaling factor from Gabor response magnitudes to match the measured neural responses to different stimulus conditions, based on their orientation tuning (a (o, b); analogous to c (o, b)) in units of ΔF/F (Fig. 12b, c). We termed these scaled Gabor responses the “Tuning only model” as they are solely based on the scaled, orientation-dependent Gabor filter responses at the central stimulus location (Fig. 12b, c; Fig. S2). The tuning-only model responses were computed as follows:

$$tuning\ only\ responses(o,b) = w_t \times c(o,b)$$

(10)

The “Tuning only model” fit showed the strongest deviation from the experimental data in the narrow condition. Since narrow stimuli also have a more regular spatial structure compared to ORI stimuli (Figure 14c), we hypothesized that the difference between the modeled and experimental data could be explained by surround modulation. We, therefore, included an additional interaction between the center and spatially surrounding Gabor responses within

the 30° by 30° patch into the model. This "Suppression model" used divisive surround suppression to normalize center responses:

$$suppression\ model\ responses(o,b) = \frac{1}{n_F}\sum_{f\ \epsilon\ F}^{n_F}\frac{w_c \times m(o,f,b,x_c,y_c)}{1 + w_s \times \max(0, s(o,f,b) + t)}$$

(11)

Here, w_c, w_s, and t represent fitted model parameters, and s was defined as the average response magnitude of Gabor filters with the same orientation and spatial frequency tuning as in the center but from the surround region.

$$s(o,f,b) = \frac{1}{|P|}\sum_{x,y\epsilon P}^{n=|P|} m(o,f,b,x,y)$$

(12)

With:

$$P = \{(x,y)|x \in X, y \in Y and\ d(x,y) > lim_i\ and\ d(x,y) < lim_o\}$$

$$d(x,y) = \sqrt{(x-x_c)^2 + (y-y_c)^2}$$

(13)

Here, P defines the set of "valid" 2D (x,y) coordinates within the 37 x 37 grid of Gabor responses that define the surround region. |P| denotes the cardinality of the set P. P depends on the model parameters lim_i and lim_o, which define the minimum and maximum distance from the center Gabor response to qualify as part of the surround (converted from degrees to corresponding units).

We then fitted the suppression model using these five parameters: the minimum and maximum distance from the center at which responses were considered as contributing to the surround modulation (lim_i, lim_o), the weight of the center and surround responses (w_c, w_s) and the surround response threshold (t). This surround suppression acted on the level of individual frame responses and was fit to the response over the entire stimulus dataset using the fminsearchbnd-function (John D'Errico, 2006; Matlab file-exchange). This allowed us to limit the surround region from 7.5° to 15° displacement from the central Gabor, as well as limit w_c and w_s to positive values. For the presented model, we used the parameters of the best fit over 100 individual model fits with random parameter initializations. The presented, best fitting model used the following parameters: lim_i= 11.19°, lim_o=12.19°, w_c=11.9, w_s=1295.1 and t=-0.2 arbitrary units. The resulting orientation-tuning dependent fit is shown in Fig. 12b, c. From these fitted narrow and ORI responses, we then derived the "response modulation" shown in Fig. 12d.

To generate a fit for the expected recruitment, based on the “tuning only model” as well as the “suppression model”, we used a simple linear regression without an intercept term to translate the fitted ORI response modulation to orientation-dependent neuronal recruitment (Fig. 12e). Our underlying assumption was that a higher response amplitude should be reflected in a higher fraction of neurons reaching the response criterion, resulting in larger recruitment. The fitted recruitment is derived from the Gabor response magnitude:

$$recruitment\ modulation(o) = \beta \times response\ modulation(o) \tag{14}$$

The fit for translating response to recruitment modulation was $\beta_{\text{tuning only model}}$=0.134 for the Tuning only model and $\beta_{\text{suppression model}}$=1.141 for the Suppression model. Lastly, we computed the mean-corrected coefficient of determination (R^2) to evaluate how well our models described orientation-dependent differences in the observed broadband response and recruitment modulation.

The processed data has been made publicly available in a shard folder in the Zenodo platform [https://doi.org/10.5281/zenodo.14605878] and can be freely downloaded from there.

The publicly shard natural image library (Tkačik et al., 2011) that we have used for natural feature distribution analysis in Fig. 7a and Fig. 9a is accessible at

[http://tofu.psych.upenn.edu/~upennidb].

A source data excel file that allows one to plot all quantitative figures of this thesis and look at more statistics details is provided with the publication:

[https://doi.org/10.1038/s41467-025-58003-1]

Scripts that allow one to do stimulus analysis, data analysis as well as the Gabor filter bank models can be found in GitHub:

[https://github.com/AachenBrainLab/Eyes_Wide_Bandwidth].

REFERENCES

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

Albrecht, D. G., Farrar, S. B., & Hamilton, D. B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. *The Journal of Physiology*, *347*(1), 713–739. https://doi.org/10.1113/jphysiol.1984.sp015092

Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., Van Langen, J., & Kievit, R. A. (2021). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, *4*, 63. https://doi.org/10.12688/wellcomeopenres.15191.2

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. https://doi.org/10.1016/j.neuron.2011.11.013

Antolík, J., Hofer, S. B., Bednar, J. A., & Mrsic-Flogel, T. D. (2016). Model Constrained by Visual Hierarchy Improves Prediction of Neural Responses to Natural Scenes. *PLOS Computational Biology*, *12*(6), e1004927. https://doi.org/10.1371/journal.pcbi.1004927

Atick, J. J., & Redlich, A. N. (1992). What Does the Retina Know about Natural Scenes? *Neural Computation*, *4*(2), 196–210. https://doi.org/10.1162/neco.1992.4.2.196

Baddeley, R., Abbott, L. F., Booth, M. C. A., Sengpiel, F., Freeman, T., Wakeman, E. A., & Rolls, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *264*(1389), 1775–1783. https://doi.org/10.1098/rspb.1997.0246

Bányai, M., Lazar, A., Klein, L., Klon-Lipok, J., Stippinger, M., Singer, W., & Orbán, G. (2019). Stimulus complexity shapes response correlations in primary visual cortex. *Proceedings of the National Academy of Sciences*, *116*(7), 2723–2732. https://doi.org/10.1073/pnas.1816766116

Barbera, D., Priebe, N. J., & Glickfeld, L. L. (2022). Feedforward mechanisms of cross-orientation interactions in mouse V1. *Neuron*, *110*(2), 297-311.e4. https://doi.org/10.1016/j.neuron.2021.10.017

Barlow, H. B. (1961). Possible Principles Underlying the Transformations of Sensory Messages. In W. A. Rosenblith (Ed.), *Sensory Communication* (pp. 216–234). The MIT Press. https://doi.org/10.7551/mitpress/9780262518420.003.0013

REFERENCES

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

Bergen, J. R., & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-13*(5), 857–863. https://doi.org/10.1109/TSMC.1983.6313080

Berkes, P., Orbán, G., Lengyel, M., & Fiser, J. (2011). Spontaneous Cortical Activity Reveals Hallmarks of an Optimal Internal Model of the Environment. *Science*, *331*(6013), 83–87. https://doi.org/10.1126/science.1195870

Bharioke, A., Munz, M., Brignall, A., Kosche, G., Eizinger, M. F., Ledergerber, N., Hillier, D., Gross-Scherf, B., Conzelmann, K.-K., Macé, E., & Roska, B. (2022). General anesthesia globally synchronizes activity selectively in layer 5 cortical pyramidal neurons. *Neuron*, *110*(12), 2024-2040.e10. https://doi.org/10.1016/j.neuron.2022.03.032

Bolaños, F., Orlandi, J. G., Aoki, R., Jagadeesh, A. V., Gardner, J. L., & Benucci, A. (2024). Efficient coding of natural images in the mouse visual cortex. *Nature Communications*, *15*(1), 2466. https://doi.org/10.1038/s41467-024-45919-3

Bonin, V., Histed, M. H., Yurgenson, S., & Reid, R. C. (2011). Local Diversity and Fine-Scale Organization of Receptive Fields in Mouse Visual Cortex. *The Journal of Neuroscience*, *31*(50), 18506–18521. https://doi.org/10.1523/JNEUROSCI.2974-11.2011

Bowers, J. S., Malhotra, G., Dujmović, M., Montero, M. L., Tsvetkov, C., Biscione, V., Puebla, G., Adolfi, F. G., Hummel, J., Heaton, R. F., Evans, B. D., Mitchell, J., & Blything, R. (2022). *Deep Problems with Neural Network Models of Human Vision*. https://doi.org/10.31234/osf.io/5zf4s

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Buccino, A. P., Hurwitz, C. L., Garcia, S., Magland, J., Siegle, J. H., Hurwitz, R., & Hennig, M. H. (2020). SpikeInterface, a unified framework for spike sorting. *eLife*, *9*, e61834. https://doi.org/10.7554/eLife.61834

Burkhalter, A., D'Souza, R. D., Ji, W., & Meier, A. M. (2023). Integration of Feedforward and Feedback Information Streams in the Modular Architecture of Mouse Visual Cortex. *Annual Review of Neuroscience*, *46*(1), 259–280. https://doi.org/10.1146/annurev-neuro-083122-021241

REFERENCES

Busse, L., Wade, A. R., & Carandini, M. (2009). Representation of Concurrent Stimuli by Population Activity in Visual Cortex. *Neuron*, *64*(6), 931–942. https://doi.org/10.1016/j.neuron.2009.11.004

Cadena, S. A., Denfield, G. H., Walker, E. Y., Gatys, L. A., Tolias, A. S., Bethge, M., & Ecker, A. S. (2019). Deep convolutional models improve predictions of macaque V1 responses to natural images. *PLOS Computational Biology*, *15*(4), e1006897. https://doi.org/10.1371/journal.pcbi.1006897

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62. https://doi.org/10.1038/nrn3136

Carandini, M., Heeger, D. J., & Anthony Movshon, J. (1999). Linearity and Gain Control in V1 Simple Cells. In P. S. Ulinski, E. G. Jones, & A. Peters (Eds.), *Cerebral Cortex* (Vol. 13, pp. 401–443). Springer US. https://doi.org/10.1007/978-1-4615-4903-1_7

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002). Nature and Interaction of Signals From the Receptive Field Center and Surround in Macaque V1 Neurons. *Journal of Neurophysiology*, *88*(5), 2530–2546. https://doi.org/10.1152/jn.00692.2001

Chen, X., Han, F., Poo, M., & Dan, Y. (2007). Excitatory and suppressive receptive field subunits in awake monkey primary visual cortex (V1). *Proceedings of the National Academy of Sciences*, *104*(48), 19120–19125. https://doi.org/10.1073/pnas.0706938104

Clark, D. A., Fitzgerald, J. E., Ales, J. M., Gohl, D. M., Silies, M. A., Norcia, A. M., & Clandinin, T. R. (2014). Flies and humans share a motion estimation strategy that exploits natural scene statistics. *Nature Neuroscience*, *17*(2), 296–303. https://doi.org/10.1038/nn.3600

Cody, P. A., & Tzounopoulos, T. (2022). Neuromodulatory Mechanisms Underlying Contrast Gain Control in Mouse Auditory Cortex. *The Journal of Neuroscience*, *42*(28), 5564–5579. https://doi.org/10.1523/JNEUROSCI.2054-21.2022

Coen-Cagli, R., Kohn, A., & Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. *Nature Neuroscience*, *18*(11), 1648–1655. https://doi.org/10.1038/nn.4128

Couto, J., Musall, S., Sun, X. R., Khanal, A., Gluf, S., Saxena, S., Kinsella, I., Abe, T., Cunningham, J. P., Paninski, L., & Churchland, A. K. (2021). Chronic, cortex-wide imaging of specific cell populations during behavior. *Nature Protocols*, *16*(7), 3241–3263. https://doi.org/10.1038/s41596-021-00527-z

Created in BioRender. Balla, E. (2025). [Dataset]. https://BioRender.com/g05l789.

REFERENCES

Dan, Y., Atick, J. J., & Reid, R. C. (1996). Efficient Coding of Natural Scenes in the Lateral Geniculate Nucleus: Experimental Test of a Computational Theory. *The Journal of Neuroscience*, *16*(10), 3351–3362. https://doi.org/10.1523/JNEUROSCI.16-10-03351.1996

Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, *20*(10), 847–856. https://doi.org/10.1016/0042-6989(80)90065-6

David, S. V., Vinje, W. E., & Gallant, J. L. (2004). Natural Stimulus Statistics Alter the Receptive Field Structure of V1 Neurons. *The Journal of Neuroscience*, *24*(31), 6991–7006. https://doi.org/10.1523/JNEUROSCI.1422-04.2004

DeAngelis, G. C., Robson, J. G., Ohzawa, I., & Freeman, R. D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex. *Journal of Neurophysiology*, *68*(1), 144–163. https://doi.org/10.1152/jn.1992.68.1.144

Di Santo, S., Dipoppa, M., Keller, A., Roth, M., Scanziani, M., & Miller, K. D. (2022). *The combination of feedforward and feedback processing accounts for contextual effects in visual cortex*. https://doi.org/10.1101/2022.05.27.493753

Ding, Z., Tran, D. T., Ponder, K., Cobos, E., Ding, Z., Fahey, P. G., Wang, E., Muhammad, T., Fu, J., Cadena, S. A., Papadopoulos, S., Patel, S., Franke, K., Reimer, J., Sinz, F. H., Ecker, A. S., Pitkow, X., & Tolias, A. S. (2023). *Bipartite invariance in mouse primary visual cortex*. https://doi.org/10.1101/2023.03.15.532836

Dyballa, L., Hoseini, M. S., Dadarlat, M. C., Zucker, S. W., & Stryker, M. P. (2018). Flow stimuli reveal ecologically appropriate responses in mouse visual cortex. *Proceedings of the National Academy of Sciences*, *115*(44), 11304–11309. https://doi.org/10.1073/pnas.1811265115

Felsen, G., & Dan, Y. (2005). A natural approach to studying vision. *Nature Neuroscience*, *8*(12), 1643–1646. https://doi.org/10.1038/nn1608

Felsen, G., Touryan, J., Han, F., & Dan, Y. (2005). Cortical Sensitivity to Visual Features in Natural Scenes. *PLoS Biology*, *3*(10), e342. https://doi.org/10.1371/journal.pbio.0030342

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, *4*(12), 2379. https://doi.org/10.1364/JOSAA.4.002379

REFERENCES

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, *33*(2), 173–193. https://doi.org/10.1016/0042-6989(93)90156-Q

Fiser, A., Mahringer, D., Oyibo, H. K., Petersen, A. V., Leinweber, M., & Keller, G. B. (2016). Experience-dependent spatial expectations in mouse visual cortex. *Nature Neuroscience*, *19*(12), 1658–1664. https://doi.org/10.1038/nn.4385

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

Freeman, J., Ziemba, C. M., Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (2013). A functional and perceptual signature of the second visual area in primates. *Nature Neuroscience*, *16*(7), 974–981. https://doi.org/10.1038/nn.3402

Freeman, T. C. B., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppression without Inhibition in Visual Cortex. *Neuron*, *35*(4), 759–771. https://doi.org/10.1016/S0896-6273(02)00819-X

Froudarakis, E., Berens, P., Ecker, A. S., Cotton, R. J., Sinz, F. H., Yatsenko, D., Saggau, P., Bethge, M., & Tolias, A. S. (2014). Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nature Neuroscience*, *17*(6), 851–857. https://doi.org/10.1038/nn.3707

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/10.1016/j.celrep.2024.114639

Fu, J., Shrinivasan, S., Ponder, K., Muhammad, T., Ding, Z., Wang, E., Ding, Z., Tran, D. T., Fahey, P. G., Papadopoulos, S., Patel, S., Reimer, J., Ecker, A. S., Pitkow, X., Haefner, R. M., Sinz, F. H., Franke, K., & Tolias, A. S. (2023). *Pattern completion and disruption characterize contextual modulation in mouse visual cortex* [Preprint]. Neuroscience. http://biorxiv.org/lookup/doi/10.1101/2023.03.13.532473

Gallimore, C. G., Ricci, D., & Hamm, J. P. (2023). *Spatiotemporal dynamics across visual cortical laminae support a predictive coding framework for interpreting mismatch responses* [Preprint]. Neuroscience. http://biorxiv.org/lookup/doi/10.1101/2023.04.17.537173

Ganguli, D., & Simoncelli, E. P. (2014). *Implicit encoding of prior probabilities in optimal neural populations*.

REFERENCES

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), 12587–12600. https://doi.org/10.1523/JNEUROSCI.1124-14.2014

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, *14*(7), 926–932. https://doi.org/10.1038/nn.2831

Goris, R. L. T., Simoncelli, E. P., & Movshon, J. A. (2015). Origin and Function of Tuning Diversity in Macaque Visual Cortex. *Neuron*, *88*(4), 819–831. https://doi.org/10.1016/j.neuron.2015.10.009

Guan, S.-C., Zhang, S.-H., Zhang, Y.-C., Tang, S.-M., & Yu, C. (2020). Plaid Detectors in Macaque V1 Revealed by Two-Photon Calcium Imaging. *Current Biology*, *30*(5), 934-940.e3. https://doi.org/10.1016/j.cub.2020.01.005

Guizar-Sicairos, M., Thurman, S. T., & Fienup, J. R. (2008). Efficient subpixel image registration algorithms. *Optics Letters*, *33*(2), 156. https://doi.org/10.1364/OL.33.000156

Guo, K., Robertson, R. G., Mahmoodi, S., & Young, M. P. (2005). Centre-surround interactions in response to natural scene stimulation in the primary visual cortex. *European Journal of Neuroscience*, *21*(2), 536–548. https://doi.org/10.1111/j.1460-9568.2005.03858.x

Guo, Z. V., Li, N., Huber, D., Ophir, E., Gutnisky, D., Ting, J. T., Feng, G., & Svoboda, K. (2014). Flow of Cortical Activity Underlying a Tactile Decision in Mice. *Neuron*, *81*(1), 179–194. https://doi.org/10.1016/j.neuron.2013.10.020

Haider, B., Krause, M. R., Duque, A., Yu, Y., Touryan, J., Mazer, J. A., & McCormick, D. A. (2010). Synaptic and Network Mechanisms of Sparse and Reliable Visual Cortical Activity during Nonclassical Receptive Field Stimulation. *Neuron*, *65*(1), 107–121. https://doi.org/10.1016/j.neuron.2009.12.005

Han, X., Vermaercke, B., & Bonin, V. (2022). Diversity of spatiotemporal coding reveals specialized visual processing streams in the mouse cortex. *Nature Communications*, *13*(1), 3249. https://doi.org/10.1038/s41467-022-29656-z

Harris, K. D., & Mrsic-Flogel, T. D. (2013). Cortical connectivity and sensory coding. *Nature*, *503*(7474), 51–58. https://doi.org/10.1038/nature12654

Hasani, H., Soleymani, M., & Aghajan, H. (2019). Surround Modulation: A Bio-inspired Connectivity Structure for Convolutional Neural Networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 32). Curran

Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2019/file/c535e3a7f97daf1c4b1eb03cc8e31623-Paper.pdf

Hermundstad, A. M., Briguglio, J. J., Conte, M. M., Victor, J. D., Balasubramanian, V., & Tkačik, G. (2014). Variance predicts salience in central sensory processing. *eLife*, *3*, e03722. https://doi.org/10.7554/eLife.03722

Hill, D. N., Mehta, S. B., & Kleinfeld, D. (2011). Quality Metrics to Accompany Spike Sorting of Extracellular Signals. *Journal of Neuroscience*, *31*(24), 8699–8705. https://doi.org/10.1523/JNEUROSCI.0971-11.2011

Hubel, D. H., & Wiesel, T. N. (1959). *Receptive Fields of single neurons in the cat's striate cortex.* 148(3):574-91.

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

Jones, H. E., Grieve, K. L., Wang, W., & Sillito, A. M. (2001). Surround Suppression in Primate V1. *Journal of Neurophysiology*, *86*(4), 2011–2028. https://doi.org/10.1152/jn.2001.86.4.2011

Juavinett, A. L., & Callaway, E. M. (2015). Pattern and Component Motion Responses in Mouse Visual Cortical Areas. *Current Biology*, *25*(13), 1759–1764. https://doi.org/10.1016/j.cub.2015.05.028

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydın, Ç., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., … Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, *551*(7679), 232–236. https://doi.org/10.1038/nature24636

Kalatsky, V. A., & Stryker, M. P. (2003). New Paradigm for Optical Imaging. *Neuron*, *38*(4), 529–545. https://doi.org/10.1016/S0896-6273(03)00286-1

Kaschube, M. (2014). Neural maps versus salt-and-pepper organization in visual cortex. *Current Opinion in Neurobiology*, *24*, 95–102. https://doi.org/10.1016/j.conb.2013.08.017

Katzner, S., Busse, L., & Carandini, M. (2011). GABA Inhibition Controls Response Gain in Visual Cortex. *The Journal of Neuroscience*, *31*(16), 5931–5941. https://doi.org/10.1523/JNEUROSCI.5753-10.2011

REFERENCES

Kayser, C. (2004). Processing of complex stimuli and natural scenes in the visual cortex. *Current Opinion in Neurobiology*, *14*(4), 468–473. https://doi.org/10.1016/j.conb.2004.06.002

Keller, A. J., Dipoppa, M., Roth, M. M., Caudill, M. S., Ingrosso, A., Miller, K. D., & Scanziani, M. (2020). A Disinhibitory Circuit for Contextual Modulation in Primary Visual Cortex. *Neuron*, *108*(6), 1181-1193.e8. https://doi.org/10.1016/j.neuron.2020.11.013

Keller, A. J., Houlton, R., Kampa, B. M., Lesica, N. A., Mrsic-Flogel, T. D., Keller, G. B., & Helmchen, F. (2017). Stimulus relevance modulates contrast adaptation in visual cortex. *eLife*, *6*, e21589. https://doi.org/10.7554/eLife.21589

Keller, A. J., Roth, M. M., & Scanziani, M. (2020). Feedback generates a second receptive field in neurons of the visual cortex. *Nature*, *582*(7813), 545–549. https://doi.org/10.1038/s41586-020-2319-4

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. *Neuron*, *100*(2), 424–435. https://doi.org/10.1016/j.neuron.2018.10.003

Kirchberger, L., Mukherjee, S., Schnabel, U. H., Van Beest, E. H., Barsegyan, A., Levelt, C. N., Heimel, J. A., Lorteije, J. A. M., Van Der Togt, C., Self, M. W., & Roelfsema, P. R. (2021). The essential role of recurrent processing for figure-ground perception in mice. *Science Advances*, *7*(27), eabe1833. https://doi.org/10.1126/sciadv.abe1833

Kirchberger, L., Mukherjee, S., Self, M. W., & Roelfsema, P. R. (2023). Contextual drive of neuronal responses in mouse V1 in the absence of feedforward input. *Science Advances*, *9*(3), eadd2498. https://doi.org/10.1126/sciadv.add2498

Kohler, P. J., Clarke, A., Yakovleva, A., Liu, Y., & Norcia, A. M. (2016). Representation of Maximally Regular Textures in Human Visual Cortex. *The Journal of Neuroscience*, *36*(3), 714–729. https://doi.org/10.1523/JNEUROSCI.2962-15.2016

Kreile, A. K., Bonhoeffer, T., & Hübener, M. (2011). Altered Visual Experience Induces Instructive Changes of Orientation Preference in Mouse Visual Cortex. *The Journal of Neuroscience*, *31*(39), 13911–13920. https://doi.org/10.1523/JNEUROSCI.2143-11.2011

Ladret, H. J., Cortes, N., Ikan, L., Chavane, F., Casanova, C., & Perrinet, L. U. (2023). Cortical recurrence supports resilience to sensory variance in the primary visual cortex. *Communications Biology*, *6*(1), 667. https://doi.org/10.1038/s42003-023-05042-3

REFERENCES

Lamme, V. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of Neuroscience*, *15*(2), 1605–1615. https://doi.org/10.1523/JNEUROSCI.15-02-01605.1995

Lazar, A., Klein, L., Klon-Lipok, J., Bányai, M., Orbán, G., & Singer, W. (2024). Paying attention to natural scenes in area V1. *iScience*, *27*(2), 108816. https://doi.org/10.1016/j.isci.2024.108816

Lee, H., Tanabe, S., Wang, S., & Hudetz, A. G. (2021). Differential Effect of Anesthesia on Visual Cortex Neurons with Diverse Population Coupling. *Neuroscience*, *458*, 108–119. https://doi.org/10.1016/j.neuroscience.2020.11.043

Leon, P. S., Vanzetta, I., Masson, G. S., & Perrinet, L. U. (2012). Motion clouds: Model-based stimulus synthesis of natural-like random textures for the study of motion perception. *Journal of Neurophysiology*, *107*(11), 3217–3226. https://doi.org/10.1152/jn.00737.2011

Lerner, T. N., Shilyansky, C., Davidson, T. J., Evans, K. E., Beier, K. T., Zalocusky, K. A., Crow, A. K., Malenka, R. C., Luo, L., Tomer, R., & Deisseroth, K. (2015). Intact-Brain Analyses Reveal Distinct Information Carried by SNc Dopamine Subcircuits. *Cell*, *162*(3), 635–647. https://doi.org/10.1016/j.cell.2015.07.014

Liska, J. P., Rowley, D. P., Nguyen, T. T. K., Muthmann, J.-O., Butts, D. A., Yates, J., & Huk, A. C. (2024). Running modulates primate and rodent visual cortex differently. *eLife*, *12*, RP87736. https://doi.org/10.7554/eLife.87736

Liu, Z.-Q. (1991). Scale space approach to directional analysis of images. *Applied Optics*, *30*(11), 1369. https://doi.org/10.1364/AO.30.001369

Marques, T., Summers, M. T., Fioreze, G., Fridman, M., Dias, R. F., Feller, M. B., & Petreanu, L. (2018). A Role for Mouse Primary Visual Cortex in Motion Perception. *Current Biology*, *28*(11), 1703-1713.e6. https://doi.org/10.1016/j.cub.2018.04.012

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. https://doi.org/10.1016/j.neuron.2011.12.004

Martin, K. A. C., & Schröder, S. (2013). Functional Heterogeneity in Neighboring Neurons of Cat Primary Visual Cortex in Response to Both Artificial and Natural Stimuli. *The Journal of Neuroscience*, *33*(17), 7325–7344. https://doi.org/10.1523/JNEUROSCI.4071-12.2013

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, *21*(9), 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

REFERENCES

Matsui, T., Hashimoto, T., Murakami, T., Uemura, M., Kikuta, K., Kato, T., & Ohki, K. (2024). Orthogonalization of spontaneous and stimulus-driven activity by hierarchical neocortical areal network in primates. *Nature Communications*, *15*(1), 10055. https://doi.org/10.1038/s41467-024-54322-x

McWhinnie, H. J., Segall, M., Campbell, D., & Herskovits, M. (1970). The Influence of Culture on Visual Perception. *Art Education*, *23*(6), 30. https://doi.org/10.2307/3191488

Miao, H.-Y., & Tong, F. (2024). Convolutional neural network models applied to neuronal responses in macaque V1 reveal limited nonlinear processing. *Journal of Vision*, *24*(6), 1. https://doi.org/10.1167/jov.24.6.1

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Receptive field organization of complex cells in the cat's striate cortex. *The Journal of Physiology*, *283*(1), 79–99. https://doi.org/10.1113/jphysiol.1978.sp012489

Muir, D. R., Molina-Luna, P., Roth, M. M., Helmchen, F., & Kampa, B. M. (2017). Specific excitatory connectivity for feature integration in mouse primary visual cortex. *PLOS Computational Biology*, *13*(12), e1005888. https://doi.org/10.1371/journal.pcbi.1005888

Muir, D. R., Roth, M. M., Helmchen, F., & Kampa, B. M. (2015). Model-based analysis of pattern motion processing in mouse primary visual cortex. *Frontiers in Neural Circuits*, *9*. https://doi.org/10.3389/fncir.2015.00038

Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S., & Churchland, A. K. (2019). Single-trial neural dynamics are dominated by richly varied movements. *Nature Neuroscience*, *22*(10), 1677–1686. https://doi.org/10.1038/s41593-019-0502-4

Niell, C. M. (2015). Cell Types, Circuits, and Receptive Fields in the Mouse Visual Cortex. *Annual Review of Neuroscience*, *38*(1), 413–431. https://doi.org/10.1146/annurev-neuro-071714-033807

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

O'Hashi, K., Miyashita, M., & Tanaka, S. (2007). Experience-dependent orientation plasticity in the visual cortex of rats chronically exposed to a single orientation. *Neuroscience Research*, *58*(1), 86–90. https://doi.org/10.1016/j.neures.2007.01.005

REFERENCES

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. *Nature*, *433*(7026), 597–603. https://doi.org/10.1038/nature03274

Okazawa, G., Tajima, S., & Komatsu, H. (2015). Image statistics underlying natural texture selectivity of neurons in macaque V4. *Proceedings of the National Academy of Sciences*, *112*(4). https://doi.org/10.1073/pnas.1415146112

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*(6583), 607–609. https://doi.org/10.1038/381607a0

Palagina, G., Meyer, J. F., & Smirnakis, S. M. (2017). Complex Visual Motion Representation in Mouse Area V1. *The Journal of Neuroscience*, *37*(1), 164–183. https://doi.org/10.1523/JNEUROSCI.0997-16.2017

Papale, P., Zuiderbaan, W., Teeuwen, R. R. M., Gilhuis, A., Self, M. W., Roelfsema, P. R., & Dumoulin, S. O. (2024). V1 neurons are tuned to perceptual borders in natural scenes. *Proceedings of the National Academy of Sciences*, *121*(46), e2221623121. https://doi.org/10.1073/pnas.2221623121

Pecka, M., Han, Y., Sader, E., & Mrsic-Flogel, T. D. (2014). Experience-Dependent Specialization of Receptive Field Surround for Selective Coding of Natural Scenes. *Neuron*, *84*(2), 457–469. https://doi.org/10.1016/j.neuron.2014.09.010

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D., & Carandini, M. (2021). Striatal activity topographically reflects cortical activity. *Nature*, *591*(7850), 420–425. https://doi.org/10.1038/s41586-020-03166-8

Pologruto, T. A., Sabatini, B. L., & Svoboda, K. (2003). ScanImage: Flexible software for operating laser scanning microscopes. *BioMedical Engineering OnLine*, *2*(1), 13. https://doi.org/10.1186/1475-925X-2-13

Poort, J., & Meyer, A. F. (2021). Vision: Depth perception in climbing mice. *Current Biology*, *31*(10), R486–R488. https://doi.org/10.1016/j.cub.2021.03.066

Popović, M., Stacy, A. K., Kang, M., Nanu, R., Oettgen, C. E., Wise, D. L., Fiser, J., & Van Hooser, S. D. (2018). Development of Cross-Orientation Suppression and Size Tuning and the Role of Experience. *The Journal of Neuroscience*, *38*(11), 2656–2670. https://doi.org/10.1523/JNEUROSCI.2886-17.2018

REFERENCES

Portilla, J., & Simoncelli, E. P. (2000). A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients. *International Journal of Computer Vision*, *40*, 49–70. https://doi.org/10.1023/A:1026553619983

Priebe, N. J. (2016). Mechanisms of Orientation Selectivity in the Primary Visual Cortex. *Annual Review of Vision Science*, *2*(1), 85–107. https://doi.org/10.1146/annurev-vision-111815-114456

Purves, D., Morgenstern, Y., & Wojtach, W. T. (2015). Perception and Reality: Why a Wholly Empirical Paradigm is Needed to Understand Vision. *Frontiers in Systems Neuroscience*, *9*. https://doi.org/10.3389/fnsys.2015.00156

Rao, R. P. N., & 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

Ratzlaff, E. H., & Grinvald, A. (1991). A tandem-lens epifluorescence macroscope: Hundred-fold brightness advantage for wide-field imaging. *Journal of Neuroscience Methods*, *36*(2–3), 127–137. https://doi.org/10.1016/0165-0270(91)90038-2

Rikhye, R. V., & Sur, M. (2015). Spatial Correlations in Natural Scenes Modulate Response Reliability in Mouse Visual Cortex. *The Journal of Neuroscience*, *35*(43), 14661–14680. https://doi.org/10.1523/JNEUROSCI.1660-15.2015

Ringach, D. L. (2019). The geometry of masking in neural populations. *Nature Communications*, *10*(1), 4879. https://doi.org/10.1038/s41467-019-12881-4

Rochefort, N. L., Narushima, M., Grienberger, C., Marandi, N., Hill, D. N., & Konnerth, A. (2011). Development of Direction Selectivity in Mouse Cortical Neurons. *Neuron*, *71*(3), 425–432. https://doi.org/10.1016/j.neuron.2011.06.013

Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature*, *395*(6700), 376–381. https://doi.org/10.1038/26475

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

Roy, A., Christie, I. K., Escobar, G. M., Osik, J. J., Popović, M., Ritter, N. J., Stacy, A. K., Wang, S., Fiser, J., Miller, P., & Van Hooser, S. D. (2018). Does experience provide a permissive or instructive influence

on the development of direction selectivity in visual cortex? *Neural Development*, *13*(1), 16. https://doi.org/10.1186/s13064-018-0113-x

Rucci, M., & Poletti, M. (2015). Control and Functions of Fixational Eye Movements. *Annual Review of Vision Science*, *1*(1), 499–518. https://doi.org/10.1146/annurev-vision-082114-035742

Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters*, *73*(6), 814–817. https://doi.org/10.1103/PhysRevLett.73.814

Rust, N. C., & Movshon, J. A. (2005). In praise of artifice. *Nature Neuroscience*, *8*(12), 1647–1650. https://doi.org/10.1038/nn1606

Rust, N. C., Schwartz, O., Movshon, J. A., & Simoncelli, E. P. (2005). Spatiotemporal Elements of Macaque V1 Receptive Fields. *Neuron*, *46*(6), 945–956. https://doi.org/10.1016/j.neuron.2005.05.021

Sakai, K., & Tanaka, S. (2000). Spatial pooling in the second-order spatial structure of cortical complex cells. *Vision Research*, *40*(7), 855–871. https://doi.org/10.1016/S0042-6989(99)00230-8

Samonds, J. M., Feese, B. D., Lee, T. S., & Kuhlman, S. J. (2017). Nonuniform surround suppression of visual responses in mouse V1. *Journal of Neurophysiology*, *118*(6), 3282–3292. https://doi.org/10.1152/jn.00172.2017

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience*, *4*(8), 819–825. https://doi.org/10.1038/90526

Segal, I. Y., Giladi, C., Gedalin, M., Rucci, M., Ben-Tov, M., Kushinsky, Y., Mokeichev, A., & Segev, R. (2015). Decorrelation of retinal response to natural scenes by fixational eye movements. *Proceedings of the National Academy of Sciences*, *112*(10), 3110–3115. https://doi.org/10.1073/pnas.1412059112

Self, M. W., Lorteije, J. A. M., Vangeneugden, J., Van Beest, E. H., Grigore, M. E., Levelt, C. N., Heimel, J. A., & Roelfsema, P. R. (2014). Orientation-Tuned Surround Suppression in Mouse Visual Cortex. *Journal of Neuroscience*, *34*(28), 9290–9304. https://doi.org/10.1523/JNEUROSCI.5051-13.2014

Sengpiel, F., Sen, A., & Blakemore, C. (1997). Characteristics of surround inhibition in cat area 17: *Experimental Brain Research*, *116*(2), 216–228. https://doi.org/10.1007/PL00005751

Sengpiel, F., Stawinski, P., & Bonhoeffer, T. (1999). Influence of experience on orientation maps in cat visual cortex. *Nature Neuroscience*, *2*(8), 727–732. https://doi.org/10.1038/11192

REFERENCES

Sharpee, T. O., Sugihara, H., Kurgansky, A. V., Rebrik, S. P., Stryker, M. P., & Miller, K. D. (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature*, *439*(7079), 936–942. https://doi.org/10.1038/nature04519

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences*, *98*(4), 1935–1940. https://doi.org/10.1073/pnas.98.4.1935

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural Image Statistics and Neural Representation. *Annual Review of Neuroscience*, *24*(1), 1193–1216. https://doi.org/10.1146/annurev.neuro.24.1.1193

Simoncini, C., Perrinet, L. U., Montagnini, A., Mamassian, P., & Masson, G. S. (2012). More is not always better: Adaptive gain control explains dissociation between perception and action. *Nature Neuroscience*, *15*(11), 1596–1603. https://doi.org/10.1038/nn.3229

Sit, Y. F., Chen, Y., Geisler, W. S., Miikkulainen, R., & Seidemann, E. (2009). Complex Dynamics of V1 Population Responses Explained by a Simple Gain-Control Model. *Neuron*, *64*(6), 943–956. https://doi.org/10.1016/j.neuron.2009.08.041

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*(7–8), 1078–1086. https://doi.org/10.1016/0042-6989(91)90033-2

Smith, A. T., & Edgar, G. K. (1990). The influence of spatial frequency on perceived temporal frequency and perceived speed. *Vision Research*, *30*(10), 1467–1474. https://doi.org/10.1016/0042-6989(90)90027-I

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

Stansbury, D. E., Naselaris, T., & Gallant, J. L. (2013). Natural Scene Statistics Account for the Representation of Scene Categories in Human Visual Cortex. *Neuron*, *79*(5), 1025–1034. https://doi.org/10.1016/j.neuron.2013.06.034

Stephens, G. J., Mora, T., Tkačik, G., & Bialek, W. (2013). Statistical Thermodynamics of Natural Images. *Physical Review Letters*, *110*(1), 018701. https://doi.org/10.1103/PhysRevLett.110.018701

Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M., & Harris, K. D. (2019). High-dimensional geometry of population responses in visual cortex. *Nature*, *571*(7765), 361–365. https://doi.org/10.1038/s41586-019-1346-5

REFERENCES

Susan, S. (2024). Neuroscientific insights about computer vision models: A concise review. *Biological Cybernetics*, *118*(5–6), 331–348. https://doi.org/10.1007/s00422-024-00998-9

Swindale, N. V. (1996). The development of topography in the visual cortex: A review of models. *Network: Computation in Neural Systems*, *7*(2), 161–247. https://doi.org/10.1088/0954-898X_7_2_002

Tanaka, S., Tani, T., Ribot, J., & Yamazaki, T. (2007). Chronically mountable goggles for persistent exposure to single orientation. *Journal of Neuroscience Methods*, *160*(2), 206–214. https://doi.org/10.1016/j.jneumeth.2006.09.004

Tang, S., Zhang, Y., Li, Z., Li, M., Liu, F., Jiang, H., & Lee, T. S. (2018). Large-scale two-photon imaging revealed super-sparse population codes in the V1 superficial layer of awake monkeys. *eLife*, *7*, e33370. https://doi.org/10.7554/eLife.33370

Tesileanu, T., Piasini, E., & Balasubramanian, V. (2022). Efficient processing of natural scenes in visual cortex. *Frontiers in Cellular Neuroscience*, *16*, 1006703. https://doi.org/10.3389/fncel.2022.1006703

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences*, *106*(11), 4567–4570. https://doi.org/10.1073/pnas.0811155106

Tkačik, G., Garrigan, P., Ratliff, C., Milčinski, G., Klein, J. M., Seyfarth, L. H., Sterling, P., Brainard, D. H., & Balasubramanian, V. (2011). Natural Images from the Birthplace of the Human Eye. *PLoS ONE*, *6*(6), e20409. https://doi.org/10.1371/journal.pone.0020409

Tkačik, G., Prentice, J. S., Victor, J. D., & Balasubramanian, V. (2010). Local statistics in natural scenes predict the saliency of synthetic textures. *Proceedings of the National Academy of Sciences*, *107*(42), 18149–18154. https://doi.org/10.1073/pnas.0914916107

Touryan, J., Felsen, G., & Dan, Y. (2005). Spatial Structure of Complex Cell Receptive Fields Measured with Natural Images. *Neuron*, *45*(5), 781–791. https://doi.org/10.1016/j.neuron.2005.01.029

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Uran, C., Peter, A., Lazar, A., Barnes, W., Klon-Lipok, J., Shapcott, K. A., Roese, R., Fries, P., Singer, W., & Vinck, M. (2022). Predictive coding of natural images by V1 firing rates and rhythmic synchronization. *Neuron*, *110*(7), 1240-1257.e8. https://doi.org/10.1016/j.neuron.2022.01.002

Vaiceliunaite, A., Erisken, S., Franzen, F., Katzner, S., & Busse, L. (2013). Spatial integration in mouse primary visual cortex. *Journal of Neurophysiology*, *110*(4), 964–972. https://doi.org/10.1152/jn.00138.2013

REFERENCES

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. *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. *Journal of Comparative Neurology*, *518*(11), 2051–2070. https://doi.org/10.1002/cne.22321

Vangeneugden, J., Van Beest, E. H., 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. *Current Biology*, *29*(24), 4268-4275.e7. https://doi.org/10.1016/j.cub.2019.10.037

Victor, J. D., Conte, M. M., & Chubb, C. F. (2017). Textures as Probes of Visual Processing. *Annual Review of Vision Science*, *3*(1), 275–296. https://doi.org/10.1146/annurev-vision-102016-061316

Vinje, W. E., & Gallant, J. L. (2000). Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science*, *287*(5456), 1273–1276. https://doi.org/10.1126/science.287.5456.1273

Vinje, W. E., & Gallant, J. L. (2002). Natural Stimulation of the Nonclassical Receptive Field Increases Information Transmission Efficiency in V1. *The Journal of Neuroscience*, *22*(7), 2904–2915. https://doi.org/10.1523/JNEUROSCI.22-07-02904.2002

Vogels, R., & Orban, G. (1990). How well do response changes of striate neurons signal differences in orientation: A study in the discriminating monkey. *The Journal of Neuroscience*, *10*(11), 3543–3558. https://doi.org/10.1523/JNEUROSCI.10-11-03543.1990

Walker, E. Y., Sinz, F. H., Cobos, E., Muhammad, T., Froudarakis, E., Fahey, P. G., Ecker, A. S., Reimer, J., Pitkow, X., & Tolias, A. S. (2019). Inception loops discover what excites neurons most using deep predictive models. *Nature Neuroscience*, *22*(12), 2060–2065. https://doi.org/10.1038/s41593-019-0517-x

Wang, E. Y., Fahey, P. G., Ding, Z., Papadopoulos, S., Ponder, K., Weis, M. A., Chang, A., Muhammad, T., Patel, S., Ding, Z., Tran, D., Fu, J., Schneider-Mizell, C. M., MICrONS Consortium, Da Costa, N. M., Reid, R. C., Collman, F., Da Costa, N. M., Franke, K., … Tolias, A. S. (2025). Foundation model of neural activity predicts response to new stimulus types. *Nature*, *640*(8058), 470–477. https://doi.org/10.1038/s41586-025-08829-y

REFERENCES

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), 1905–1918. https://doi.org/10.1523/JNEUROSCI.3488-10.2011

White, B. J., Berg, D. J., Kan, J. Y., Marino, R. A., Itti, L., & Munoz, D. P. (2017). Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video. *Nature Communications*, *8*(1), 14263. https://doi.org/10.1038/ncomms14263

Whiteway, M. R., Socha, K., Bonin, V., & Butts, D. A. (2019). *Characterizing the nonlinear structure of shared variability in cortical neuron populations using latent variable models* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1904.10441

Wiesbrock, C., Musall, S., & Kampa, B. M. (2022). A flexible Python-based touchscreen chamber for operant conditioning reveals improved visual perception of cardinal orientations in mice. *Frontiers in Cellular Neuroscience*, *16*, 866109. https://doi.org/10.3389/fncel.2022.866109

Williford, J. R., & Von Der Heydt, R. (2016). Figure-Ground Organization in Visual Cortex for Natural Scenes. *Eneuro*, *3*(6), ENEURO.0127-16.2016. https://doi.org/10.1523/ENEURO.0127-16.2016

Yen, S.-C., Baker, J., & Gray, C. M. (2007). Heterogeneity in the Responses of Adjacent Neurons to Natural Stimuli in Cat Striate Cortex. *Journal of Neurophysiology*, *97*(2), 1326–1341. https://doi.org/10.1152/jn.00747.2006

Yoshida, T., & Ohki, K. (2020). Natural images are reliably represented by sparse and variable populations of neurons in visual cortex. *Nature Communications*, *11*(1), 872. https://doi.org/10.1038/s41467-020-14645-x

Yu, Y., Schmid, A. M., & Victor, J. D. (2015). Visual processing of informative multipoint correlations arises primarily in V2. *eLife*, *4*, e06604. https://doi.org/10.7554/eLife.06604

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

Zhuang, J., Ng, L., Williams, D., Valley, M., Li, Y., Garrett, M., & Waters, J. (2017). An extended retinotopic map of mouse cortex. *eLife*, *6*, e18372. https://doi.org/10.7554/eLife.18372

REFERENCES

Ziemba, C. M., Freeman, J., Movshon, J. A., & Simoncelli, E. P. (2016). Selectivity and tolerance for visual texture in macaque V2. *Proceedings of the National Academy of Sciences*, *113*(22). https://doi.org/10.1073/pnas.1510847113

Ziemba, C. M., Goris, R. L. T., Stine, G. M., Perez, R. K., Simoncelli, E. P., & Movshon, J. A. (2024). Neuronal and Behavioral Responses to Naturalistic Texture Images in Macaque Monkeys. *The Journal of Neuroscience*, *44*(42), e0349242024. https://doi.org/10.1523/JNEUROSCI.0349-24.2024

Ziemba, C. M., & Simoncelli, E. P. (2021). Opposing effects of selectivity and invariance in peripheral vision. *Nature Communications*, *12*(1), 4597. https://doi.org/10.1038/s41467-021-24880-5

Zipf, G. K. (1950). Human behavior and the principle of least effort. *Journal of Clinical Psychology*, *6*(3), 306–306. https://doi.org/10.1002/1097-4679(195007)6:3<306::AID-JCLP2270060331>3.0.CO;2-7

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

APPENDIX

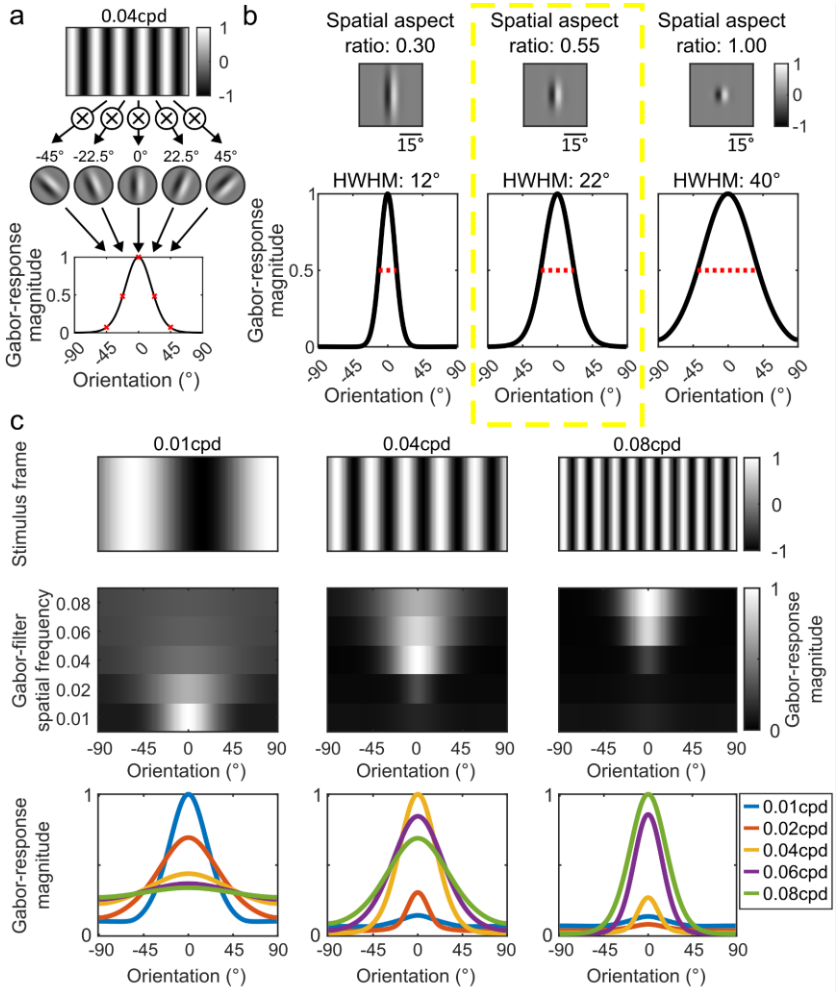


Figure S1. Gabor filter bank model to represent mouse V1 responses.

(**a**) Schematic of Gabor filter responses. Top: Example image of a sine wave grating stimulus with a spatial frequency of 0.04 cpd. Middle: different orientations of Gabor filters with a spatial frequency of 0.04 cpd and a spatial aspect ratio of 0.55, visualized for an example phase of the complex spatial kernel. Bottom: grating response curve over Gabor filters with different orientations. (**b**) The spatial aspect ratio of the used Gabor filters was adjusted to match the orientation tuning of mouse V1 neurons in response to the same 0.04cpd grating shown in a). Top row: visualization of the spatial filter kernels with different aspect ratios. The scale bar indicates the size of each Gabor filter in visual angles relative to stimulus space. Bottom row: Normalized response magnitude across different orientations of the Gabor filter with the corresponding spatial aspect ratio (title in the top row). The red line indicates the half-maximum response at which the half-width half-maximum (HWHM in the panel titles) of the resulting orientation tuning curve was calculated. The yellow box indicates the spatial aspect ratio that was used for all the following model fits. (**c**) Normalized responses of Gabor filters with different spatial frequencies and orientations to different grating stimuli. To represent the spatial frequency tuning of V1 neurons, Gabor filters for five different spatial frequencies (0.01 cpd, 0.02 cpd, 0.04 cpd, 0.06 cpd and 0.08 cpd) were used to model V1 responses. Responses for each of the Gabor filters were normalized to their peak response to a sine wave grating that matched their corresponding spatial frequency. Top row: 3 example sine wave grating stimuli with different spatial frequencies, indicated in the panel titles. Middle rows: Response magnitudes to the stimulus shown above across different orientations of the five types of Gabor filters with different spatial frequency tuning. Gabor filters displayed the strongest responses to stimuli that matched their spatial frequency tuning. Bottom: Same as middle panels but showing the tuning curves for each Gabor filter type individually for better comparability across spatial frequency tuning of the different Gabor filters. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

APPENDIX

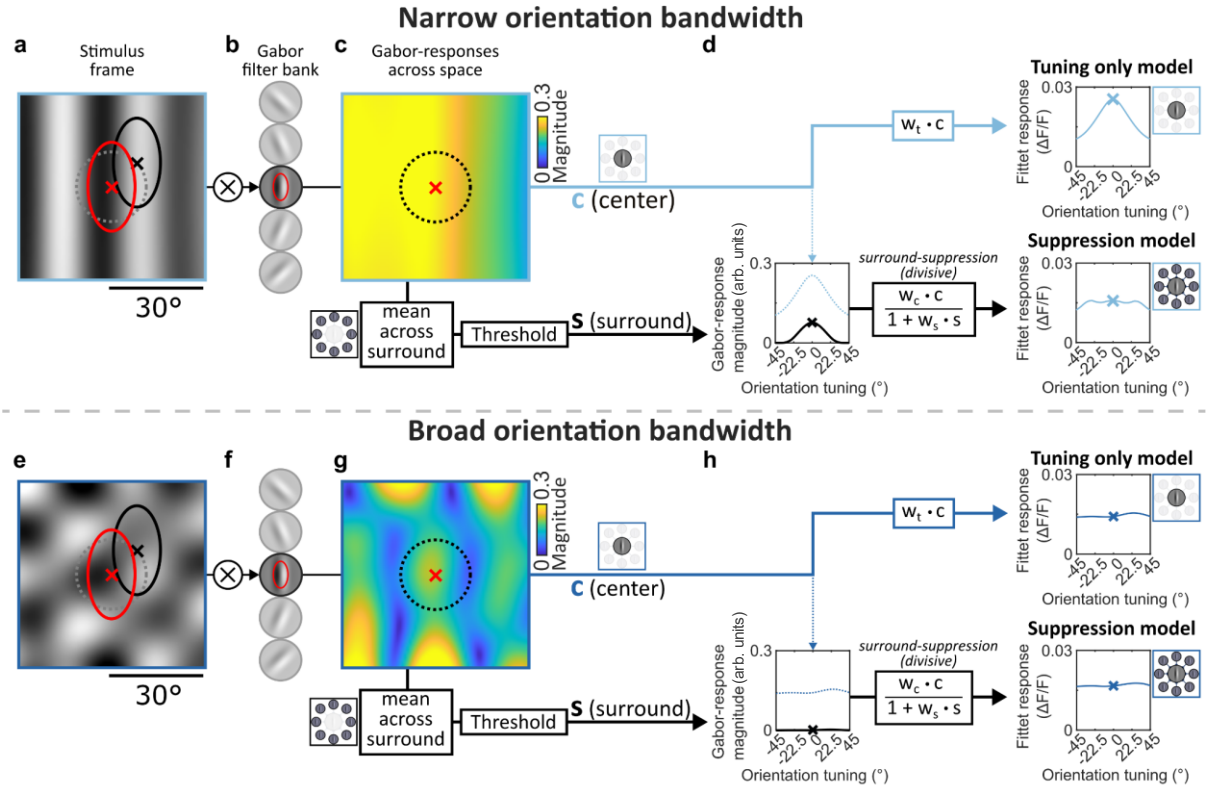


Figure S2. Fitting Gabo model responses to narrow and broad orientations to the mouse V1 responses.

(**a**) Example narrow-band frame. Red ellipse and cross indicate the outline and midpoint of the center Gabor, respectively. Black ellipse and cross indicate the location of an example Gabor filter, contributing to the surround suppression in the Suppression model. Dashed circle indicates the midpoint locations of other surround Gabor filters (radius = ~11.7°, matching the dashed line in c). (**b**) Bank of Gabor filters, comprising filters with five orientations between -45° and 45° and five different spatial frequencies (0.01, 0.02, 0.04, 0.06, 0.08 cpd). Highlighted is an example Gabor filter kernel with 0° and 0.04 cpd tuning. Red ellipse shows the half-maximum of the Gabor envelope, similar to a) (**c**) Response magnitudes of Gabor responses across space after applying the example filter of stimulus frame. Each pixel indicates the response of this Gabor at different locations in the frame (mapped in correspondence to the Gabor filter midpoint). From these responses across space, either only the center Gabor response (red cross; Tuning only model) or both the center and the surrounding responses were used (dashed circle; Suppression model). (**d**) Tuning-only model, predicts neuronal responses to the center by fitting scaling factor w_t, which scales average Gabor response magnitudes to average V1 responses (see Figure 3b, c). Tuning-only model, therefore, presents the best fit when describing neural responses as simple Gabor filters with different orientation and spatial frequency tuning, averaged over all stimulus frames of the corresponding condition. Suppression model extracts Gabor responses of the same orientation that spatially surround the center Gabor. Surround magnitudes are then averaged and thresholded with a fitted offset to obtain the surround response. The surround region limits, thresholding offset, and the center and surround weights (w_c, w_s) are simultaneously fitted and divided the center by the surround response. Tuning curve markers highlights the link to the 0° Gabor shown in (b) and (c). (**e-h**) Same models were applied to broadband stimuli and fitted simultaneously to the measured neural responses to narrow and broadband stimuli for fit optimization (see Figure 3b, c). Panels c, d, g, and h were created in BioRender. Balla, E. https://BioRender.com/g05l789 (2025). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

APPENDIX

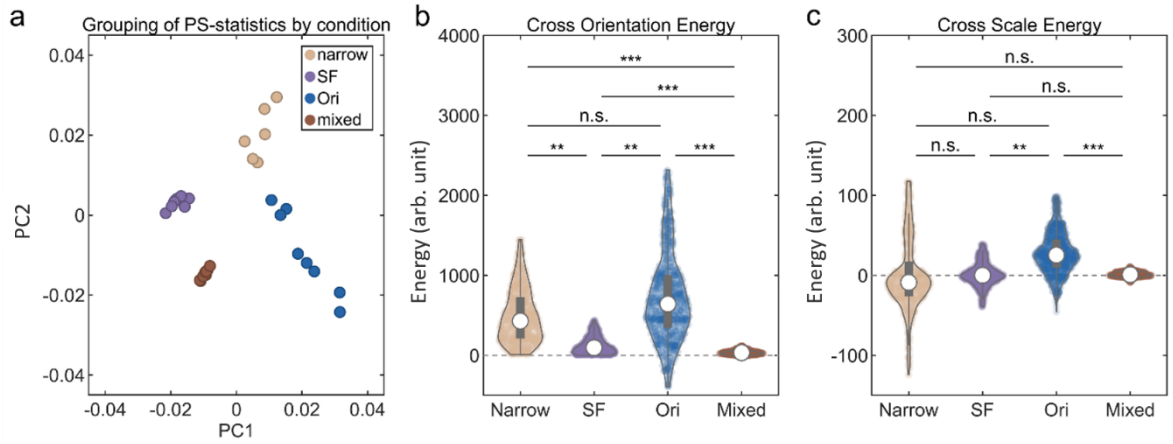


Figure S3. Stimulus separability estimated by coefficient magnitude statistics.

(**a**) Overview of overall variability and separability among the ten renderings for each stimulus condition ('Narrow' - beige, 'SF' - purple, 'ORI' - blue, 'Mixed'- brown). Texture statistics for each stimulus frame was calculated using the Portilla and Simoncelli analysis ('PS-statistics', see also Methods: 'Stimulus predictability' and 'Stimulus separability')[2]. Then, a principal component analysis across all PS-statistics (including pixel statistics, auto-correlations, magnitude statistics, as well as statistics of the low- and high-frequency residuals) was computed and averaged across all frames of a rendering. Projection onto the first two principal components (PC) is shown for individual renderings. Data points spread shows the variability in texture statistics between renderings of the same stimulus condition. Largest variability is seen in narrow and ORI conditions. However, individual stimulus renderings formed non-overlapping clusters, demonstrating that they contain distinct combinations of PS-statistics that are unambiguously distinguishable between stimulus categories. (**b**) The effect of expanding spatial frequency or orientation bandwidth on complex structures or second-order textures was examined by comparing coefficient magnitude statistics. Cross Orientation Energies, derived from PS-statistics, were computed from energy correlations between the 0° and the ±45° sub-band magnitudes at the coarsest spatial scale of the steerable image pyramid decomposition. Expanding spatial frequency bandwidth, especially in "Mixed", reduced cross-orientation energy, while orientation bandwidth had no significant effect. This suggests that expanding spatial frequency bandwidth reduces image structures (e.g., edges, bars, corners) on the coarsest spatial scale. (**c**) Comparison of cross-scale energies between stimulus conditions. Shown are average cross-correlation of 0° filter sub-band magnitudes between the two coarsest spatial scales of the image pyramid. Expanding orientation bandwidth resulted in significantly higher cross-scale energies compared to SF and Mixed. However, neither expanding spatial frequency- nor orientation bandwidth significantly changed cross-scale energies compared to narrow stimuli, suggesting that there are no strong and consistent changes in cross-scale energy. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Significance was tested with a Wilcoxon–Mann–Whitney U test with Bonferroni-correction ($\alpha = 0.0167$, 3 tests) comparing the cross energies of individual stimulus realizations, averaged over individual frames. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

APPENDIX

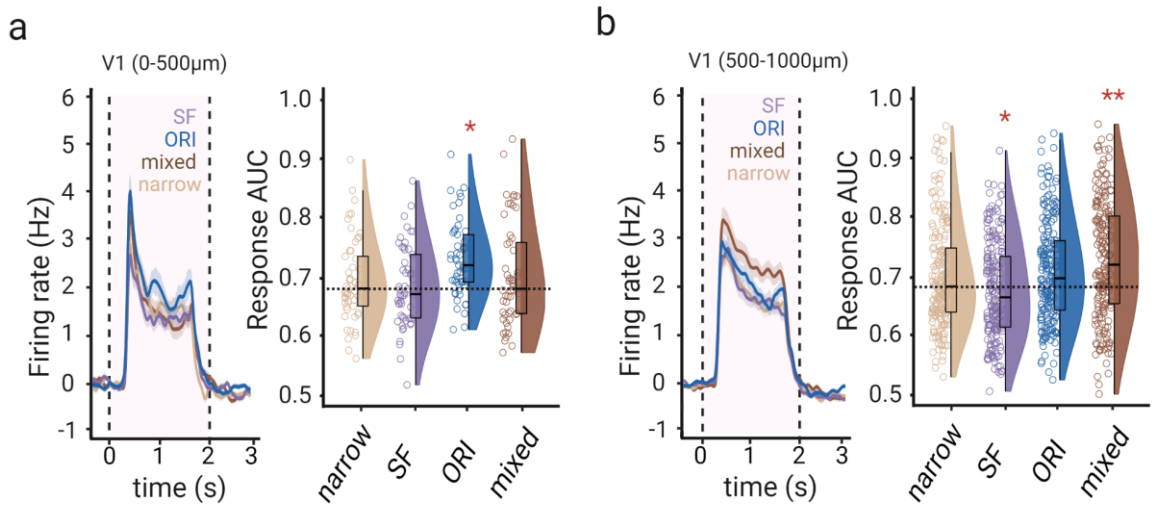


Figure S4. Broadband responses are different in the superficial and the deeper V1 neurons.

(**a**) Left: Mean firing rates of superficial (0-500 μm deep) neurons in V1. The shaded area around the solid lines shows the s.e.m. The area between the vertical dashed lines is the stimulus presentation window (n = 53 responding units from 3 animals). Right: Response AUCs against baseline for each stimulus condition. (**b**) Same as in panel a but for deeper V1 neurons between 600-1000 μm (n = 200 responding units from 3 animals). For visualization only, outliers were excluded from distributions in the panels. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Stars in all panels mark significant differences from the LME model test against narrow. The significance threshold was α = 0.0167 after Bonferroni correction for performing three tests. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

APPENDIX

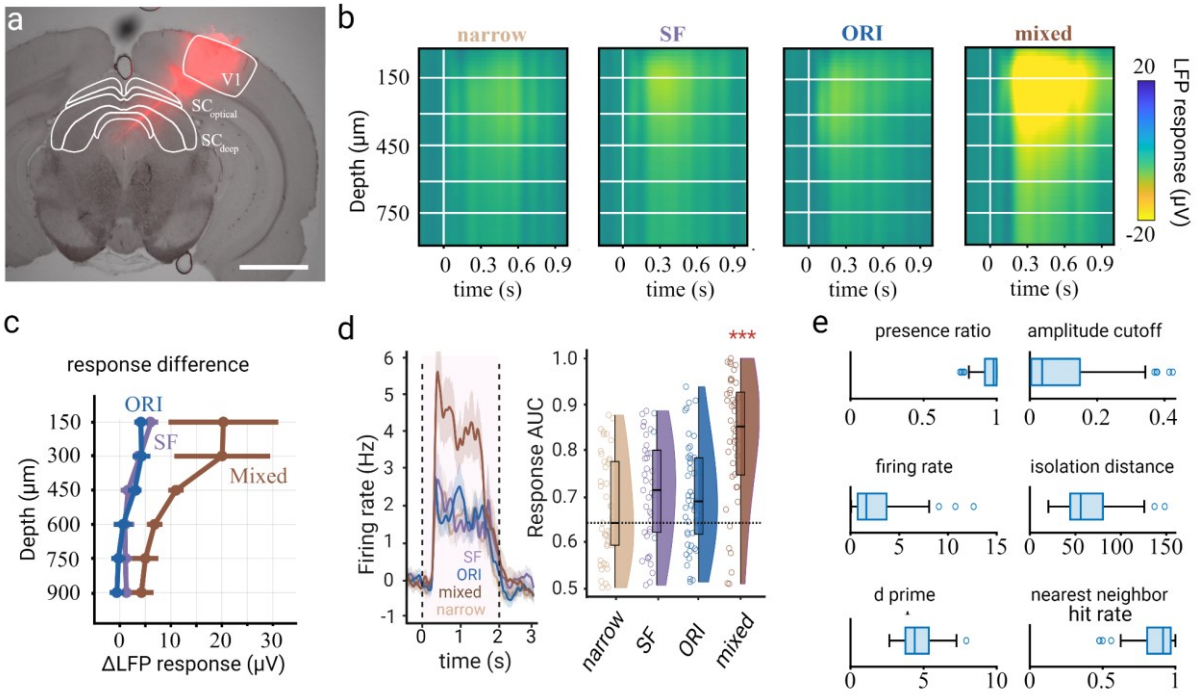


Figure S5. Neurons in the optical SC preferentially respond to the mixed broadband stimulus condition.

(**a**) Example brain slice with electrode tract, following the electrophysiological recordings. The fluorescent DiD signal from the Neuropixels probe is visible in pink, originating in V1 and reaching different SC layers. The scale bar is 1.5 mm. Brain region annotations are based on the Allen Mouse Brain Common Coordinate Framework (CCF)[3]. (**b**) Electrophysiological LFP responses to each stimulus condition at different depths of the SC. Different depths from the SC surface are marked by white lines. The SC surface was determined by the first channel with observable visual responses in the LFP, which occurred at a recording depth of 1.63±0.16 mm across six recordings from 3 mice. Colors show voltage changes after stimulus onset relative to baseline. (**c**) Differences in LFP responses for each stimulus condition (SF: purple; ORI: dark blue; mixed: brown) compared to the LFP response to the narrow stimulus across depths. Error bars show mean and s.e.m. (n = 6 recordings from 3 mice). Neural responses were strongest for the mixed condition in the superficial SC layers, and responses to all stimulus conditions were lower in deeper SC layers. (**d**) Left panel: Mean firing rate of SC neurons. The shaded area around the solid lines shows the s.e.m. The area between the vertical dashed lines is the stimulus presentation window (n = 46 responding units from 3 animals). Right panel: Response AUCs for different stimulus conditions for individual neurons. Stars mark significant differences from the LME model test against the narrow condition. The significance threshold was $\alpha = 0.0167$ after Bonferroni correction for performing three tests. For visualization only, outliers were excluded. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). (**e**) Spike sorting quality metrics for all manually selected good clusters across all Neuropixels recordings. The quality metric results are consistent with well-isolated units, such as a high presence ratio, d-prime, and nearest-neighbor hit rate, with a low amplitude cutoff. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

APPENDIX

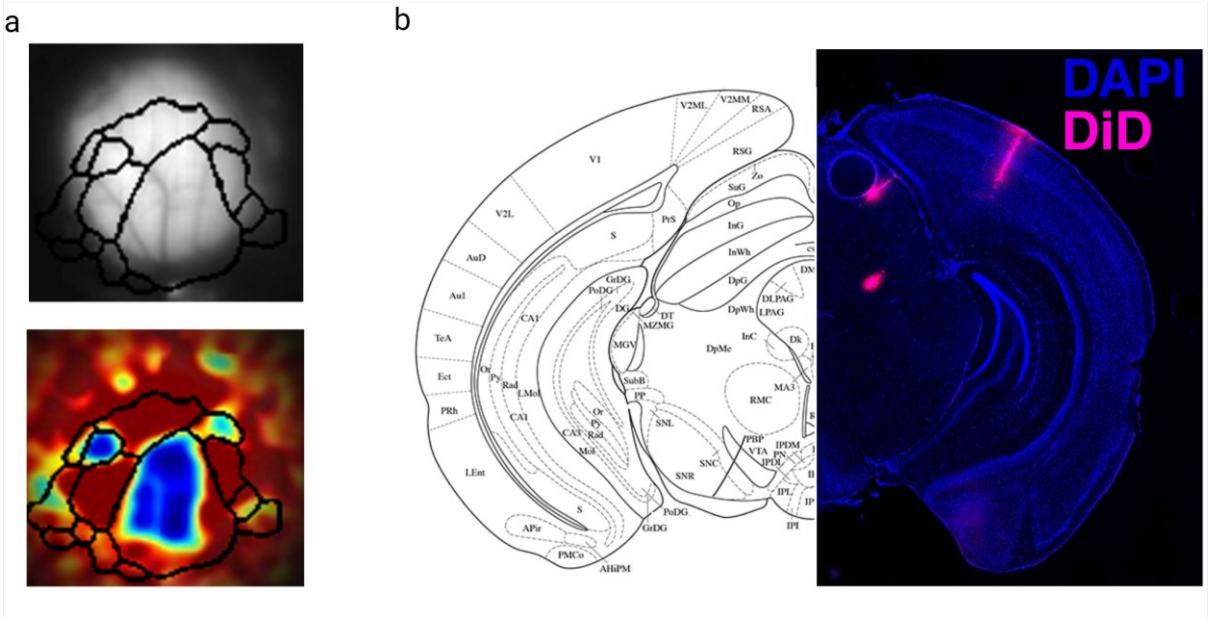


Figure S6. Retinotopic mapping of cortical visual areas and histology of the Neuropixels probe depth.

(**a**) Example retinotopic mapping result from a mouse in which we performed Neuropixels recordings. The inferred border areas are drawn on top of the vessel architecture (top) based on the visual sign map that informs visual area locations (bottom). Recordings were done in the center of V1 (large blue area in the bottom panel). (**b**) Example histology results following the electrophysiological recordings. Fluorescent marks toward the midline are from recordings where the probe reached the SC. The fluorescent trace in the cortex is from a recording where the probe was placed perpendicular to the cortical surface. Blue is DAPI staining, and pink is the DiD trace of the Neuropixels probe insertion. The opposite hemisphere is replaced by a scheme of the mouse brain atlas (The mouse brain in stereotaxic coordinates: hard cover edition, Paxinos, George, and Keith B.J. Franklin, page 60, Copyright Elsevier (2001)) at the same rostro-caudal location as the example slice. Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

APPENDIX

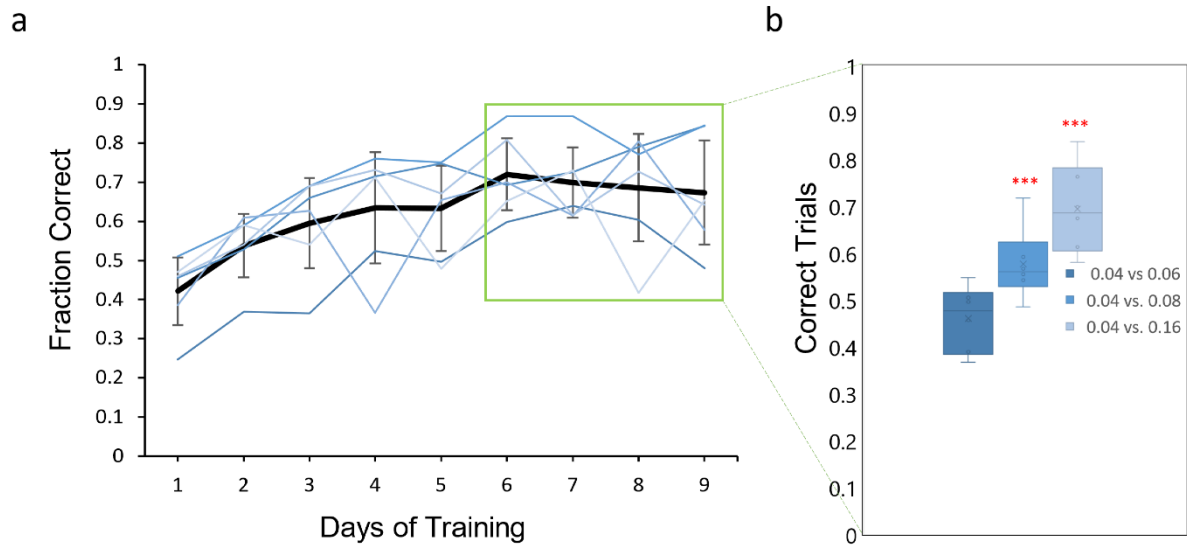


Figure S7. Learning curves of central spatial frequency discrimination in our mice.

(**a**) Shows the learning curves over nine days of training for all the six mice that took part in the central spatial frequency discrimination protocol in the touchscreen chamber. All stimuli were kept at a low SF bandwidth (0.004 cpd) and at our lowest chosen orientation bandwidth (5°), while the target stimulus had a central SF of 0.16 cpd, and the distractor stimulus had a central SF of 0.04 cpd. The thick black line at the center is the group mean. Error bars indicate the standard deviation of the mean. (**b**) To get an estimate of the central SF discrimination threshold of our mice, we let them achieve expert performance (green square containing the last four sessions in panel a), and then tested different target-distractor central SF combinations. Average discrimination performances across all six mice for various combination of target-distractor that were tested. Reducing the difference between the target and distractor had an immediate effect on discrimination performance, suggesting that the initially chosen 0.04 versus 0.16 target-distractor combination is already close to their discrimination threshold. Correct trials for '0.04 vs 0.06 cpd' = 0.46 ± 0.02, T-test against 0.5, $p = 5.9 \times 10^{-2}$; '0.04 vs 0.08 cpd' = 0.58 ± 0.02, $p = 1.4 \times 10^{-4}$; '0.04 vs 0.16 cpd' = 0.69 ± 0.02, $p = 1.4 \times 10^{-8}$, mean ± s.e.m, n = 6 mice. Significance threshold was $\alpha = 0.0167$ after Bonferroni correction for performing three tests. Box plots indicate the median (horizonal line), interquartile range (box bounds: 25th–75th percentiles), and whiskers (1.5× interquartile range). Figure is reproduced from Balla et al., Nature Communications 16, 58003 (2025), licensed under CC BY 4.0.

ACKNOWLEDGEMENTS

I would like to take a moment to appreciate everyone who contributed to this research dissertation and to my wellbeing while it was taking place.

To my Doktorvater Björn, thank you for trusting me with your projects and for giving me executive freedom in my experimental work, which though at times is harder, does indeed nurture confidence and initiative. Thank you for being open to all my far-fetched ideas even when I only had “concepts of a plan”. And generally, for sharing with us all the cool events and places to visit and your extensive repertoire of dad jokes and science anecdotes.

To my Doktoronkel Simon, thank you for jumping into this project without hesitation, for always being there in dire times and for continuously challenging me out of my comfort zone vis-à-vis scientific communication, coding and soldering. Certainly, learned a lot from your presence.

To Gerion, our lab ‘wunderkind’, thank you for being the human version of a multimodal neuron in the lab, for all the tech support, for the shared bureaucratic loathing, and for always answering my questions in a very smart but very convoluted way.

To Chris and Jenice, thank you for seeing the initial potential on the motion cloud stimulus which was the first domino for our very nice paper and also for all the constructive discussions.

To Severin, thank you for taking the time to do the electrophysiological recordings with us.

To Sandra, thank you for all the backstage support and taking off some of the bureaucratic heat from our shoulders.

My lab girlies from all the different eras:

The cottage core lot: Magda, Kerstin and Sophie (not a girly but Thomas belongs in this group too) for all the fun travels, arts and crafts sessions, coffee magic, deep talks and free group therapy. I have cherished every moment with you fine young ladies (and your +1s). Extending this appreciation also to my DnD group for excusing all the late show-ups and last-minute improvs.

The Femininomenon lot: Sacha, Irene, Alice, Danique, Natalja, Nilu, Anoushka, Julia and Fatemeh. I have a lot to be grateful about my experiences with you from cool playlists exchanged during a bus ride, to late night conversations about deep societal issues, to collecting iconic quotes. Special spotlight goes on Sacha and Irene for being close to me in

important life events and for sharing the pressures of thesis submission together in the last leg of our shared journey.

My dear family, thank you for all the love and support from afar, the late calls and the understanding, for being very unassuming and willing to learn and grow with me and my generation. This extends to my late grandpa who instilled curiosity in my mind and filled my childish world with magic and wonder and for always coming up with impromptu fairytales for his 4 -year-old, illiterate, story-time-addict niece (with an unfortunately excellent memory).

Finally, to my wonderful husband, thank you for being the epitome of empathy, consideration and support with your angelic vocals and a repertoire ranging from: “Hello, it’s me, I was wondering if after all these pages you’d like to eat” to “rooomance ninja”. Thank you for all the skills you have perfected just to make me happy like the optimized coffee making or mind-reading. Thank you for listening to my daily TED talks and especially the long windy ones during the very low writing phase and for unskillfully but bravely lying that you’d love me even if I were a worm. But most importantly thank you for helping me be the beta version of myself.

Eidesstattliche Erklärung

DECLARATION OF AUTHORSHIP

I, Elisabeta Balla declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research.

Hiermit erkläre ich an Eides statt / I do solemnly swear that:

1. This work was done wholly or mainly while in candidature for the doctoral degree at this faculty and university;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this university or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others or myself, this is always clearly attributed;

4. Where I have quoted from the work of others or myself, the source is always given. This thesis is entirely my own work, with the exception of such quotations;

5. I have acknowledged all major sources of assistance;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Parts of this work have been published before as.

Balla E, Nabbefeld G, Wiesbrock C, Linde J, Graff S, Musall S, Kampa B (2025). Broadband visual stimuli improve neuronal representation and sensory perception. Nat Commun 16, 2957 (2025). https://doi.org/10.1038/s41467-025-58003-1

Formal authors contributions (from the article): B.M.K., C.W., and E.B. conceived and planned the experiments. E.B. designed the 2-photon experiments, collected the neural data, and performed the analysis with input from B.M.K. and S.M. G.N. generated the stimuli, collected and analysed the intrinsic- and widefield imaging data, and built the model of V1. C.W. and J.L. designed the 2AFC task, collected the behavioural data, and performed the analysis. S.G. and S.M. collected and analysed the electrophysiological data. E.B. wrote the initial draft and generated the figures with input from B.M.K and S.M. The manuscript was revised and completed by B.M.K and S.M. All authors read and approved the final manuscript.

Eidesstattliche Erklärung

Note on my contribution to this study: I have performed all experiments, data-analysis, and text writing related to 2-photon imaging, with inputs from other authors. I have generated all the visual stimuli and stimulus presentation protocols for the imaging experiments (with initial onboarding from C.W., G.N and programming mentorship from S.M).

G.N. generated the intrinsic stimuli, collected and analysed the intrinsic- and widefield imaging data, and built the model of V1. Thus, parts of the methods sections [4.1.3, 4.2.2, 4.5.1, 4.7 and 4.8] text and supplementary figures [Fig. S1 – S3] are work from G.N. S.G. and S.M. collected and analyzed the electrophysiological data. Thus, parts of the methods sections [4.2.3, 4. 5.2] text and supplementary figures [Fig. S4 – S6] are work from S.M. and S.G. C.W. and J.L. designed the 2AFC task, collected the behavioural data, and performed the analysis. Thus, parts of the methods section [4.4] text and supplementary figure [Fig. S7] are work from C.W. and J.L.

In addition, we will submit part of this work for publication:

The literature research that went into section 1 (Introduction) of this manuscript will lead to a review paper publication. **Balla E.**, Kampa B. Natural scenes, Neural codes.

Formal authors contributions (in the future article, agreed by authors): B.K and E.B conceptualized the review and literature covered. E.B drafted the manuscript. B.K. supervised and edited the manuscript. All authors read and approved the final manuscript.

Note on my contribution to this section: I have done the literature research and initial text writing (with inputs from B.K). B.K provided close supervision, inputs on the text and direction of the review. He will also edit and finalize the manuscript for initial submission.

Date: on the 27/10/2025

Signature: Elisabeta Balla