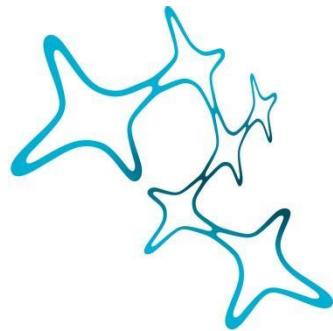

NEURAL CORRELATES OF CONTRAST NORMALIZATION IN THE *DROSOPHILA* VISUAL SYSTEM

Nadezhda Pirogova



Graduate School of
Systemic Neurosciences

LMU Munich



Dissertation at the
Graduate School of Systemic Neurosciences
Ludwig-Maximilians-Universität München

21st of December, 2022

Supervisor

Prof. Dr. Alexander Borst

Max Planck Institute of Neurobiology, Martinsried, Germany

First Reviewer:

Prof. Dr. Alexander Borst

Second Reviewer:

Prof. Dr. Laura Busse

External Reviewer

Prof. Dr. Adrienne Fairhall

Date of Submission:

21.12.2022

Date of Defense:

13.11.2023

CONTENT

1. INTRODUCTION	1
1.1 Visual Environments	1
1.1.1 Vision as One of the Primary Senses	1
1.1.2 Visual Environments and Requirements They Impose	1
1.1.3 Physics of Vision	2
1.1.4 Statistics of Visual Environments	5
1.2. Canonical Computations	6
1.2.1 Normalization is a Type of Adaptation	7
1.2.2 Effects of Adaptation	7
1.2.3 Function of Adaptation	8
1.2.4 Adaptation to Mean and Variance of the Signal	10
1.2.5 Mechanisms of Normalization	10
1.3. <i>Drosophila</i> as Model Organism in Motion Vision Research	11
1.3.1 History of <i>Drosophila</i> Research	11
1.3.2 <i>Drosophila</i> Optic Lobe	12
1.4 Tools to Study <i>Drosophila</i> Vision	14
1.4.1 Targeting Cells: The Gal4/UAS system	14
1.4.2 Visualizing Cells and Cell Activity: GFP and GCaMP	15
1.4.3 Manipulating the Circuit: TeTxLC	16
1.4.4 Recording Single Neuron Activity: Electrophysiology and Two-Photon Microscopy	17
1.4.5 Observing Natural Behavior: Behavioral Assays	18
1.4.6 Recreating the Circuit: Modeling	19
1.5 Studying Individual Elements of the Motion Detection Circuit	19
1.5.1 Computations in the <i>Drosophila</i> Optic Lobe	19
1.5.2 Motion Detection Circuit in <i>Drosophila</i> Optic Lobe	20
1.5.3 Coincidence Detectors	22
1.5.4 Beyond the First Stages of Motion Detection	24
1.6 Scopes and Aims of the Thesis	24
2. MANUSCRIPTS	26

2.1 Manuscript 1: Dynamic Signal Compression for Robust Motion Vision in Flies	26
2.2 Manuscript 2: Contrast Normalization Affects Response Time-Course of Visual Interneurons	61
3. DISCUSSION	95
3.1 Contrast Normalization in the Fly Optic Lobe	95
3.1.1 Effects of Contrast Normalization	95
3.1.2 Characteristics of Contrast Normalization	95
3.1.3 Contrast Normalization Improves Robustness of Motion Detectors	96
3.2 Mechanism of Contrast Normalization	96
3.2.1 Contrast Normalization Relies on Neural Feedback	96
3.2.2 Contrast Normalization Alters Response Time Constants	97
3.2.3 Contrast Normalization and Shunting Inhibition	98
3.2.4 State- and Stimulus-Dependent Modulation of Neural Processing Speed	99
3.3 Implications for Characterization of Visual Neurons	100
3.3.1 Stimuli to Characterize Filtering Properties of the Cell	100
3.3.2 Stimulus Design and Inconsistencies in Cell Characterization	101
3.3.3 Stimuli Profiles and Motion Detection	103
3.3.4 Synthetic Stimuli to Scan the Motion Detection Circuit	105
3.4. Conclusion and Future Directions	106
3.4.1 Searching for the Normalizing Neuron	107
3.4.2 Validating the Role of Shunting Inhibition	108
3.4.3 Recapitulating the Effects of the Dynamic Surround with Octopamine	108
4. REFERENCES	109
5. APPENDIX	119
5.1 Acknowledgments	119
5.2 List of Manuscripts and Author Contributions	120
5.3 Curriculum Vitae	121
5.4 Affidavit	123

LIST OF FIGURES

Figure 1. Physics of vision.

Figure 2. Neurons involved in motion detection in the *Drosophila* optic lobe.

Figure 3. Gal4/UAS System.

Figure 4. Visualizing Cells and Cell Activity.

Figure 5. Tools to Study *Drosophila* Vision.

Figure 6. Motion detectors.

Figure 7. Stimuli to characterize cell filtering properties.

Figure 8. Stimuli spatial profiles.

SUMMARY

The fruit fly *Drosophila melanogaster* has long become a paramount model organism for research in life sciences. As a result of the fly's high temporal resolution and its reliable optomotor response - a reflex that helps it compensate for movements of the environment, - *Drosophila* lends itself exceptionally well to the study of vision and, in particular, the mechanism of motion detection.

In the wild, *Drosophila* is active throughout the day, with especially high levels of activity at dawn and dusk, the periods of time when the visuals of the environment are changing rapidly. The fruit flies can also be found in a variety of habitats, from the expanse of an open field to the inside of a cluttered kitchen. Altogether, *Drosophila* encounters a variety of visual statistics it must employ to robustly respond to the outside world and succeed in finding food, escaping predators, and carrying out courtship behavior.

In my thesis, I focused on the effects of visual contrast, i.e., differences in brightness in the environment, on the fly's motion vision. I studied the impact of the surround contrast on the filtering properties of the visual interneurons within the motion detection circuit, including the first direction-selective T4 and T5 cells and their main inputs, and how the fly compensates for the changes in contrast to faithfully match the direction and speed of its movement to the external motion under various contrast conditions.

Firstly, in **Manuscript 1**, we established the existence of contrast normalization in the early visual system of *Drosophila* and demonstrated its suppressive effect on the response amplitude at higher contrasts. We determined where contrast normalization first arises in the optic lobe and identified the main inputs into the T4 and T5 cells that exhibit contrast normalizing properties. We comprehensively characterized the normalization process: namely, it is fast, not dependent on the direction of motion, its effect comes from outside the receptive field of a cell and increases in strength with the size of the visual surround. Additionally, we demonstrated that the normalization relies on neuronal feedback and showed that adding a contrast normalization stage to the existing models of motion detection improves their robustness, matching their performance to the results obtained in behavioral experiments.

In **Manuscript 2**, we further investigated the effects of contrast normalization on the main inputs to T4 and T5 cells, now focusing on its effect on the filtering properties of the cells. We demonstrated that spatially or temporally dynamic surrounds elicit contrast normalization, while static ones do not. We further showed that, in addition to the suppressive effect on the amplitude, contrast normalization speeds up the kinetics of the response and confirmed that

this effect is not due to signal saturation and involves a change in the filtering properties of the cell.

In summary, we elucidated the role of contrast normalization in the motion detection circuit in the early visual system of *Drosophila*, comprehensively described the characteristics of the normalization process, and outlined its effects on the filtering properties of the cells. We also emphasized the potential role of shunting inhibition and narrowed down the search for the main candidates in the contrast normalization mechanism, paving the way for future studies to further delve into the contrast normalization circuit and implementation mechanism.

1.INTRODUCTION

1.1 Visual Environments

1.1.1 Vision as One of the Primary Senses

Imagine a completely ordinary situation most encounter every day—crossing a busy road in an urban area: a wide street is full of cars whizzing by in both directions, different colors of traffic lights are blinking on and off, high-contrast black and white stripes on the road denote the pedestrian areas and street borders. Now, consider the amount of information one is getting at such a moment and the importance of the decisions that have to be made then and there and you will see this mundane occurrence in a very different light.

To survive, animals need to be able to react to the environment around them fast and in an appropriate manner, but how quick and suitable such reactions can be is limited by how well they can perceive the world through the input they receive from their senses. To make a decision about crossing the road, one needs to be able to detect motion in the environment and to correctly deduce if the source of the motion is external (a car moving) or a result of self-motion (us crossing the road). To this end, a sensor is necessary that is responsive, fast, and reliable, with respect to both transmitting the information and filtering out the noise, and it is able to carry out those tasks while using the least amount of energy possible. Such a sensor could ultimately mean the difference between life and death, increasing the selection pressure to improve the sensor's characteristics and become the fastest, most sensitive, most reliable, and requiring the least amount of energy.

1.1.2 Visual Environments and Requirements They Impose

Despite the differences in animals' habitats, vision is one of the primary senses for the majority of animals, and animals that rely on eyesight for survival are prevalent. The ability of humans and other animals to perceive their environment visually is the result of complex interactions between light, primary visual organs (commonly, eyes), and the brain. Among various species, eye design depends not only on the constraints imposed by the habitat and visual tasks the animal is likely to encounter but also on purely biological limitations that arise from the design of perceptual systems. These factors affect the ability to resolve light

patterns in the environment, and finding a compromise between them and optimizing eye design is vital for the animal's survival.

Specialized behaviors also require different information to be extracted from the environment. For example, a fast predator needs to resolve rapidly changing images to find and target its prey and might favor resolution over sensitivity, while for a slow-flying nocturnal insect, capturing as much light as possible might be beneficial, even if it comes at the cost of image resolution (Gonzalez-Bellido et al., 2011). Ultimately, as visual demands differ greatly depending on the tasks an animal commonly performs and the habitat the animal exists in, the eye design reflects the animal's lifestyle and the versatility of neural information processing in its visual system (Geisler, 2008).

1.1.3 Physics of Vision

Sensory systems are under a lot of evolutionary pressure as sensory organs are the connecting point between an organism and its environment. The evolution of eyes, ears, and noses is driven by the need to perceive and interpret electromagnetic and acoustic waves and molecule concentration, respectively, to represent the characteristics of the objects in the outside world. In vision, specifically, the ability to detect electromagnetic waves in the environment and interpret them into perceived images is highly influenced by the physics of the eye, which determines the specifics of how the world is seen. Here, the physics of vision influences which and how much information is available, as well as how fast the information is processed and transmitted for further computations.

Animal eyes show unique adaptations that arise from the requirements imposed by the environment and from unavoidable biological constraints, such as the size of the animal and its eyes, the signaling ability of the receptors and neurons, or the organic material available for the construction of the visual sensors. These morphological and physiological adaptations of visual systems include varying the shape of the eye or the position of the light detector, creating an acute zone with increased sensitivity, modifying downstream processing channels, and many more, and these adaptations create neural responses that are tuned to spatial and temporal characteristics of the visual environment specific to the animal (Gonzalez-Bellido et al., 2011).

In animals with eyes, visual processing starts when light hits photoreceptors. The main processes along the visual cascade that occur in vertebrate and invertebrate eyes are briefly summarized below:

In the **vertebrate eye** (see Figure 1A), the amount of light entering the eye is regulated by the pupil. The light is first refracted as it passes through the cornea and the lens. The cornea, the lens, and the vitreous body together act as a compound lens that projects an inverted image onto a light-sensitive membrane, the retina, where it is converted into a neural signal. The retina encompasses over 100 million specialized photoreceptive cells to gather light, namely, rods and cones. Rods are responsible for vision at low light levels: they are sensitive to a broad range of light wavelengths but low-resolution; while cones are responsible for vision at higher light levels and have one of the three photopigment types—red, green, or blue—that correspond to a particular wavelength and peak response. The process of phototransduction, i.e., the conversion of light into a change in the electrical potential of the cell membrane, starts as the light is absorbed by the photopigments. Phototransduction involves a sequence of activations of signaling proteins that cause further opening and closing of ion channels, resulting in an electrical signal in the photoreceptor cell membrane. This signal travels along the optic nerve, ultimately arriving at the occipital lobe and being transmitted further upstream, where the analyzed pattern is translated into images a conscious mind can interpret.

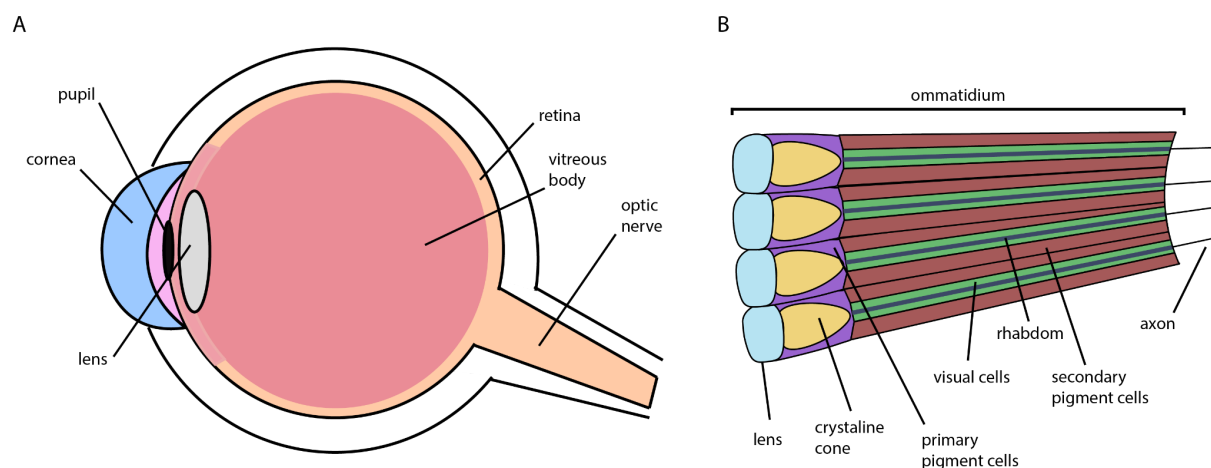


Figure 1. Physics of vision. Schematic representation of: A. a vertebrate eye, B. an invertebrate eye.

In the **compound eye** (see Figure 1B), a visual organ of arthropods, light first falls onto the ommatidia—tiny independent detector units with a corneal lens and photoreceptor cells. Each ommatidium is oriented to point in a slightly different direction and focuses light from a narrow visual field onto sensory cells and a rhabdom, housed within. Rhabdoms are created by a multitude of microvilli that extend from the sensory cells and contain high concentrations of a light-sensitive pigment rhodopsin. Additionally, each ommatidium

contains screening pigments in primary and secondary pigment cells, which stop stray light scattered from the neighboring ommatidia from entering the ommatidium. Phototransduction starts when rhodopsin absorbs light and converts it into electrical signals that are sent by the sensory cells further to the pair of optic lobes (or optic ganglia), one innervating each compound eye. The whole image perceived by an arthropod eye is, thus, built up from the individual inputs from the numerous ommatidia.

One of the main visual cues for vertebrate and invertebrate eyes in the outside environment is **optic flow**, the distribution of local motion directions and velocities across the visual field. Optic flow is what allows the animal to determine the direction of external and self-motion, guiding its course control. Local motion signals within an optic flow field are calculated downstream of the primary sensory receptors. Detecting spatial contrast requires a neural computation that compares the luminance changes at two separate points in space.

As the physics of the eye determines an animal's ability to detect light and process visual information, it also influences how animals perceive motion by impacting the perception of the visual characteristics of their environments. Some of the main characteristics of the physics of the eye are shortly summarised below.

Luminance sensitivity (the efficiency of detecting light) and **contrast sensitivity** (the efficiency of detecting differences in light intensities) correspond to the physiological operational range of vision. They are defined by light intensities normally encountered by the animal, and their lower boundary sets the absolute sensitivity threshold of the visual system (Frolov, 2016). Here, the spectral sensitivity of the photoreceptor's visual pigment determines the probability that a photon of a given wavelength (energy) will be absorbed and converted into a neural signal. The spectral sensitivity can be affected by a number of physical characteristics, such as the lens size, the rate of thermal isomerization of the photopigment, or, more specifically for flies, by the ommatidial acceptance angle, the size of the rhabdomeres, and, most importantly, the identity of the photopigment (Osorio and Vorobyev, 2005; Frolov, 2016). In comparison to fruit flies, human eyes are sensitive only to a narrow range of wavelengths and do not function well in low light.

Visual acuity or **spatial resolution** measures how much detail of the outside environment can be discerned. It is defined by the number of visual points, i.e. pixels, that make up the image and is directly related to the density of visual sensors - retinal cells in the human eye or ommatidia and the number of ommatidia in the eye of a fruit fly (Land, 1997; Sanjayan, 2018). As such, it is also generally correlated to the size of the eye (Harmening, 2017), as the larger size allows for more visual detectors. Thus, the vertebrate eye, in comparison to that of

insects, offers significant advantages, such as better depth perception and increased discrimination of the details. In addition, spatial resolution is also affected by the diffractions of the lenses that cause a pixel of an image to blur. This is especially the case in compound eyes, as diffraction increases with the shrinking of the lenses (Gonzalez-Bellido et al., 2011). Flies also do not have a specialized zone with more acute vision (fovea), so the spatial resolution stays very low, around 5° in fruit flies, throughout the whole eye.

The challenge of a low image resolution and the absence of a fovea in flies make their large **field of view**, i.e., the angle of the observable world seen at any given moment, invaluable. For the fruit flies, their field of view covers almost the entirety of the visual space, with individual ommatidia providing visual fragments that are combined into a large overall image. Unlike vertebrates with forward-facing eyes, *Drosophila* can see predators approaching from any direction at the same time, thanks to their panoramic vision.

1.1.4 Statistics of Visual Environments

According to the efficient coding hypothesis (Barlow, 1961), a theoretical model of sensory coding in the brain, sensory systems adapt in such a way as to encode and transmit the maximum amount of sensory information while minimizing the cost of the encoding and transmission. To optimize a sensory system for encoding the sensory signals found in the animal's natural environment, sensory neurons need to match their limited range of responses to the range of input signals an animal would get from a typical environment, and also to ensure that all responses occur equally often (Laughlin, 1981).

As the visual systems of animals are adapted to exploit the regularities of natural environments, the responses are adapted to and determined by the prominent statistics and fundamental components of the natural images, i.e., the inputs of the sensory system in question (Simoncelli, 2003; Geisler, 2008; Cai et al., 2020). These statistics and components of the visual environments include, among others, luminance, contrast, color, noise, and spatial range and structure of an image, all shortly described below.

At the very basic level, **luminance** is the intensity of light passing through, emitted from, or reflected from a particular area. Depending on the intensity or the amount of light, varying degrees of light, dark, or color are produced. Mean luminances vary between scenes and can locally vary within one scene. Dark and light asymmetries are found naturally in the environments, and the fact that, in an average scene, there are more dark pixels than light is also exploited by the visual systems (Chichilnisky and Kalmar, 2002). **Contrast** is the difference in luminance of the image points within the same field of view and defines the

range of luminances found within an image. As with luminance, there are large variations of local contrast in natural images. **Color** is derived from the spectrum of light interacting with the physics of the environment and objects and the characteristics of the light-perceiving cells of the eyes. Colors contribute to the contrast of a visual scene.

The amount of **noise** in a given environment determines a signal-to-noise ratio of the visual signal and influences how much visual information is needed to make a decision, i.e., it shortens or lengthens the integration time (Gold and Shadlen, 2007).

The distribution of distances or the **range** within a given image is important for understanding depth and for the perception of motion, for example, through motion parallax: an object moving at a constant speed across the visual field appears to move a greater amount if it is closer than when it is at a greater distance. The spatial pattern of luminance and color within an image determines its **spatial structure**. Here, strong features tend to cluster in natural images, as shown in the spatial auto-correlation function (or the Fourier amplitude spectrum) that stays relatively consistent across different natural scenes (Geisler, 2008).

1.2. Canonical Computations

Brains of different animals carry out the same tasks to ensure the animal's survival: they notice the motion of a predator, remember the location of the nest, detect the smell of something edible and calculate the distance to its source, and more. To this end, the brain has to perceive the characteristic of interest, process the initial information about it, and perform further computations on the same statistics of the outside world. Despite the wide diversity of habitats and early evolutionary divergence of the species, some of these computations remain remarkably similar across different animals. These are the canonical computations (Carandini and Heeger, 2011)—standard computational modules that apply the same fundamental operations, conserved across a variety of animals and in a variety of contexts. Such canonical computations can rely on diverse circuits and mechanisms, occur in different brain regions, and encompass unique neural components. These computations often stay preserved across evolutionary divergent species as a result of the adaptation that occurs in the sensory organs to compensate for the organs' limitations, such as the finite number of states available to record the outside world.

1.2.1 Normalization is a Type of Adaptation

Adaptation encompasses many processes, occurring on a variety of timescales, and is defined by an adjustment to environmental conditions, be it evolutionary adaptation in an organism's traits or a change in response to sustained stimulus characteristics. Here, I will focus on the shorter timescale of adaptation. Normalization, specifically, describes a type of scaling of a response, such as the adaptation that happens when a response to a sustained central stimulus is affected, or normalized, by a stimulus in the surround. Adaptation has been proposed as one of the canonical computations of sensory systems (Carandini and Heeger, 2011), since it has been demonstrated in a variety of model organisms, at multiple levels of processing—from the most peripheral sensory systems to decision-making—and in various sensory modalities. It has been shown to occur in response to different statistics of the environment and to be implemented in different brain regions via a wide array of mechanisms. Some examples of adaptation include light adaptation in the turtle retina (Baylor and Fuortes, 1970), adaptation to temporal contrast in salamander and mouse retinal ganglion cells (RGCs; Chander and Chichilinsky, 2001; Kim and Rieke, 2001), adaptation to sound intensity in the mechanosensory neurons (Clemens et al., 2018), or normalization in the olfactory population codes (Olsen et al., 2010) in fruit flies.

1.2.2 Effects of Adaptation

The most common effect of normalization is the suppression of the response amplitude to stronger inputs (Kim and Rieke, 2001; Baccus and Meister, 2002; Harris et al., 2000; Olsen et al., 2010). This reduction in the gain of the signal reflects the decrease in the cell sensitivity with the increase in the input strength and plays an important role in keeping the output of the system within a given range. The amplitude suppression can be seen in, for example, the suppression of graded potentials or the decrease in the frequency of spikes. This effect has been shown to occur very quickly (Kim and Rieke, 2001; Baccus and Meister, 2002; Wark et al., 2007), on the order of time expected from a normal visual fixation, and to have a prolonged effect, still altering the response when the next stimulus is presented after a 0.4 s pause (Camp et al., 2011).

In addition to amplitude suppression, in some animals, adaptation has been shown to hyperpolarize a neuron's membrane potential, reducing the baseline of the signal and making it harder for a spiking neuron to fire (Demb, 2008; Harris et al., 2000).

Normalization has also been shown to affect the temporal dynamics of the signal and change the filtering properties of a neuron to be more band-pass-like, making the response more transient or creating a phase advance in responses to sinusoidal inputs (Kim and Rieke, 2001; Chander and Chichilnisky, 2001; Baccus and Meister, 2002; Liu and Gollisch, 2015).

In vision, adaptation, and specifically normalization, constantly adjusts visual coding, altering it according to the visual stimuli that an animal is exposed to. Adaptation can occur on different time scales—from fast processes, occurring on the order of milliseconds, to slow ones lasting for hours, with both types of the process sometimes occurring in the same circuit or even in the same cells (Kim and Rieke, 2001; Baccus and Meister, 2002; Wark et al., 2007). To distinguish between the different timescales of adaptation in response to the change in contrast, some separate between the essentially instantaneous “gain control” and a slower process of “adaptation” (Baccus and Meister, 2002; Demb, 2008). By this definition, contrast gain control is a fast process that is active on the timescale of 0.1 s, occurs in response to a change in the input, reduces sensitivity, and sharpens the temporal response of visual neurons. Contrast adaptation, on the other hand, is a slower activity-dependent process that reduces sensitivity, but does not change the temporal dynamics of the response.

Paradoxically, adaptation of the signal has also been shown to sensitize responses. For example, in the inner retina of zebrafish, an increase in contrast can lead to depression of the strong and facilitation of the weak bipolar cell responses, with distinct circuits causing simultaneous increases or decreases in the responses of the same cell type (Nikolaev et al., 2013). Moreover, stimuli not tailored to evoke robust responses from the tested neurons can cause response facilitation and shifts in tuning toward the adapting stimulus itself (Wissig and Kohn, 2012).

1.2.3 Function of Adaptation

Functionally, the main task of adaptation has been suggested to be maximizing the amount of information. Barlow's efficient coding hypothesis (Barlow, 1961) states that, given a finite capacity to transmit information, neural systems employ an optimally efficient coding strategy to represent the inputs that they typically process. In this way, adaptation and, specifically, normalization have been shown to increase the efficiency of neuronal coding and increase the amount of sensory information an organism can use (Laughlin, 1989), largely as a consequence of the following effects:

- A. Adaptation equalizes the response and prevents saturation.

Adaptation in the retina shifts the operating range of photoreceptors to match the prevailing stimulus distribution and maximize sensitivity (Harris et al. 2000; Carandini and Heeger, 2011), to improve stimulus discriminability by modulating stimulus salience (Solomon et al., 2014), and to decrease integration time and gain, protecting the response from saturation and improving the retina's ability to encode fast temporal changes (Kim and Rieke, 2001; Laughlin, 1989). As a result, adaptation equalizes cell responses to different stimuli allowing the neurons to use their dynamic range uniformly and to efficiently encode the stimulus distribution (Olsen et al., 2010).

B. Adaptation reduces redundancy in the signals.

Normalization facilitates the decoding of distributed neural representations and discrimination among representations of different stimuli (Carandini and Heeger, 2011), and reduces redundancy by removing spatially and temporally redundant signal components in sensory representations (Laughlin, 1989; Solomon et al., 2014). For example, in cases when response variability is correlated between visual cortex neurons with similar orientation tuning, feedforward divisive normalization of a neuron's inputs effectively decorrelates their variability, making estimates of a stimulus parameter more accurate than those based on non-normalized inputs (Tripp, 2012). Normalization has also been shown to decorrelate the activity in different glomeruli, creating independent responses (Olsen et al., 2010). These computational benefits of normalization make it possible to accommodate more information within the signal of a single cell (Laughlin, 1989).

C. Adaptation increases robustness to noise in the responses.

Increased normalization has been shown to diminish trial-to-trial response variability, decreasing the noise in the responses (Ruff et al., 2016, Coen-Cagli and Solomon, 2019). In extreme cases of noise cancellation, the computational benefits of normalization can even include max-pooling (winner-take-all competition) (Carandini and Heeger, 2011).

Overall, adaptation reduces the effects of some fundamental extrinsic and intrinsic limitations on sensory processing. However, adaptation often involves trade-offs: sensitivity to one component of the stimulus is sacrificed for sensitivity to another. For example, adaptation can prevent saturation by trading gain for temporal acuity (Laughlin, 1989) or can provide invariance with respect to some stimulus dimensions at the expense of others (Carandini and Heeger, 2011).

1.2.4 Adaptation to Mean and Variance of the Signal

Normalization is well studied in the first sensory peripheral cells in the visual system and their adaptation to light. Photoreceptor sensitivities are modulated by the background light intensity, with the photoreceptors' response curves shifting their steepest portion to correspond to the background intensity, i.e., the mean of the input (Carandini and Heeger, 2011). Normalization to the mean, which causes photoreceptors to adjust their operating point to match mean light intensity, has been demonstrated in the retina of vertebrates (for example, in turtles by Baylor and Fuortes (1970) and in monkeys by Makman (1970)), as well as invertebrates (for example, in blowflies and dragonflies by Laughlin and Hardie (1978)).

Adaptation to the variance of the visual signal, i.e., the contrast has also been demonstrated in the early visual systems of a variety of vertebrates, such as cats (Shapley and Victor, 1979), rabbits (Brown and Masland, 2001), salamanders (Garvert and Gollisch, 2013), and mice (Khani and Gollisch, 2017). Contrast normalization in these animals reduces the sensitivity of the cells and alters the temporal profiles of the signals, making them more transient.

Contrast normalization in the *Drosophila* visual circuit, at the early stages of motion detection, was not researched in detail before. Adaptation to the changes in the statistics of the environment, however, has been shown at a later stage of the fly's motion detection circuit (Borst et al., 2005). Moreover, adaptation to the variance of the input in other sensory modalities in fruit flies has been demonstrated: for example, fast adaptation has been found to arise in mechanosensory neurons (Johnston's organ neurons) in the antenna in response to both the antennal position (the stimulus mean) and to the sound intensity (the stimulus variance), with both computations occurring in the same neuron but partly independent from each other (Clemens et al., 2018).

1.2.5 Mechanisms of Normalization

As normalization has been found in a wide range of model organisms, brain areas, and modalities, a vast array of different feedforward and feedback neural mechanisms has been suggested as a source of this adaptation, such as the presynaptic lateral inhibition (Olsen et al., 2010) that causes shunting inhibition (Carandini et al., 1997) and synaptic depression (Abbott et al., 1997; Jarsky et al., 2011).

Lateral inhibition (Olsen et al., 2010) has a clear role in gain control in visual processing. It creates a negative feedback loop that normalizes the response of the neuron by an activity of a larger “normalizing” pool. This divisive normalization scales with the total neuronal

activity and keeps the system output within a given range. Lateral inhibition can also cause the unexpected sensitization effect by adapting the normalizing pool itself, normalizing its activity instead of the activity of the neuron (Wissig and Kohn, 2012; Nikolaev et al., 2013).

In the mammalian visual circuit (Demb, 2008; Jarsky et al., 2011), a number of specific biophysical mechanisms for contrast normalization have been identified. Contrast adaptation in RGCs causes the neuron to adapt to the input and prevents response saturation and is implemented via reduced gain or membrane hyperpolarization. Here, both local synaptic processes acting on the input currents (such as synaptic depression at the bipolar-to-ganglion cell synapse), as well as ganglion cell-intrinsic mechanisms involved in the generation of spike trains (such as inactivation of sodium currents and activation of potassium currents) have been implicated in retinal contrast adaptation (Kim and Rieke, 2001; Demb, 2008; Khani and Gollisch, 2017). In the RGCs inputs, upstream rod bipolar and AII amacrine cells, increased mean or variance of presynaptic membrane potential also causes decreased gain. Here, Ca^{2+} channel inactivation and vesical depletion have been demonstrated to cause the adaptation to the mean, while adaptation to the variance has been shown to be caused by vesicle depletion alone.

The existence and the mechanisms behind contrast normalization in the early visual system of *Drosophila* are still speculative.

1.3. *Drosophila* as Model Organism in Motion Vision Research

1.3.1 History of *Drosophila* Research

Since its first introduction as a model organism in genetics, *Drosophila melanogaster* has had a long history in research in general and in circuits neuroscience in particular. Areas, where research on fruit flies has been awarded the Nobel Prize encompass genetics, olfaction, the immune system, and circadian rhythm, and the knowledge about the fly itself expanded during the years since it was first used in research.

Drosophila's developmental stages are well understood, a genetic toolbox providing cellular access has been available for decades (Rubin and Spradling, 1982), fruit fly's complete genome was sequenced and published in 2000 (Adams et al., 2000), and electron microscopy (EM) reconstruction of a complete brain of an adult *Drosophila* is now available (Zheng et al., 2018). The exhaustively developed genetic toolbox allows access to specific cell types

and single cells. Tools are available for visualization of the circuits and their activity, as well as to manipulate the circuits and disrupt their physiology.

The popularity of *Drosophila* as a model organism is additionally boosted by the unique advantages it offers, such as a short life cycle, ease of upkeep, a large number of offspring per generation, and the small size of its brain. Thus, it is not surprising that *Drosophila* has also been extensively used in research of the visual system. Though fruit flies do not have a superior spatial resolution, they rely heavily on vision for their survival and their temporal resolution allows them to perform incredibly fast aerial maneuvers.

1.3.2 *Drosophila* Optic Lobe

Drosophila has compound eyes that consist of 750-800 ommatidia, each pointing in a slightly different direction and representing a separate point in the visual space. One ommatidium houses 8 photoreceptors (R1 to R8), with 6 outer photoreceptors (R1-R6) and two inner photoreceptors (R7 and R8), located underneath. Within the photoreceptors, visual pigments are densely packed in compartments called “rhabdomeres”. Upon rhabdomere illumination, a phototransduction cascade is triggered, leading to the opening of transient receptor potential (TRP) and TRP-like (TRPL) channels that transmit the electrical signals from the retina to the next stages of visual processing.

The information is sent to the optic lobes, where it is processed by four successive neuropils: the lamina, the medulla, the lobula, and the lobula plate (see Figure 2). Between consecutive neuropils, neural fibers cross over and form the outer and the inner optic chiasma, respectively, causing the neural presentation of the visual image to be inverted in the medulla and then inverted back to the original before the lobula complex (the lobula and the lobula plate).

The lamina consists of lamina cartridges that house laminar monopolar (L) cells. The lamina receives direct input from the outer photoreceptors R1-R6 and projects further into the distal medulla (Meinertzhagen and O'Neil, 1991).

In addition to the input from the lamina (Takemura et al., 2008), the medulla receives direct input from the inner photoreceptors R7 and R8 (Fischbach and Dittrich, 1989; Takemura et al., 2017) and projects, in parallel, to the lobula and the lobula plate. The medulla itself consists of 10 layers (M1-M10) with several types of neurons (Fischbach and Dittrich, 1989), among them: Mi (medulla intrinsic) cells that connect different layers within the medulla; Tm (trans-medulla) cells that connect specific medulla layers to the lobula; TmY (trans-medulla Y) cells that connect specific layers in the medulla to the lobula and the lobula plate; and the

unique CT (complex tangential) neuron –CT1– that innervates M10 layer in the medulla and connects it to Lo1, the most proximal among the six layers (Lo1-Lo6) in the lobula.

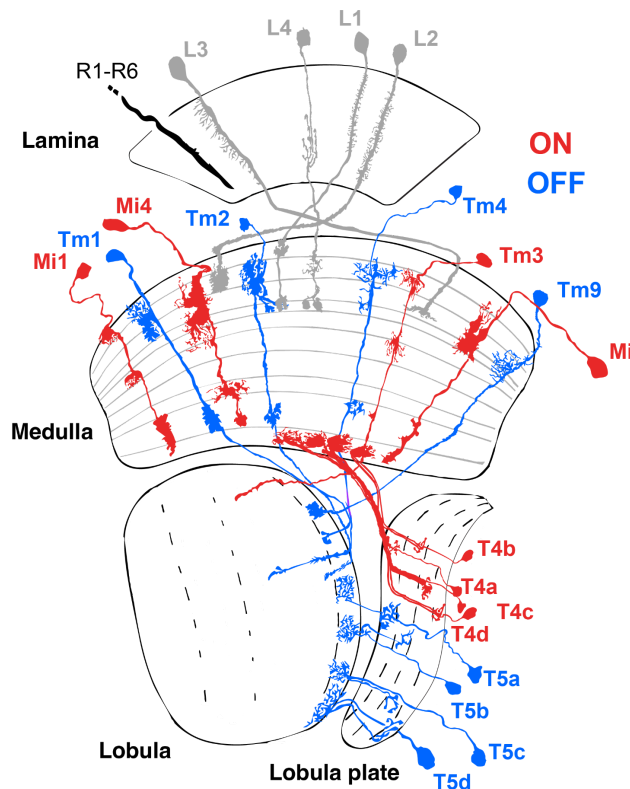


Figure 2. Neurons involved in motion detection in the *Drosophila* optic lobe. Schematic of the main elements of the motion detection circuit. Adapted from Fischbach and Dittrich (1989).

The signals received from the medulla and the lobula are processed in the four layers of the lobula plate (LoP1-LoP4) and transmitted to the central brain by visual projection neurons and various types of lobula plate tangential cells (LPTCs) (Hausen, 1984).

Each of the neuropils comprises regular columnar elements that correspond topographically to the ommatidia pattern in the retina and retain retinotopy, meaning that neighboring points in the visual space are processed by cells in the neighboring columns.

Light and dark increments are processed in the distinct parallel ON and OFF pathways in the optic lobes. The separation between these pathways is already established at the level of the lamina neurons (Joesch et al., 2010; Maisak et al., 2013). Here, the input from the outer photoreceptors R1-R6 is split depending on the polarity of the incoming signal, with the L1 cells providing input to the ON pathway that conveys information about brightness increments and the L2-L4 cells to the OFF pathway that transmits information about light decrements (Joesch et al., 2010; Takemura et al. 2011; Shinomiya et al. 2014; Meier et al. 2014).

The main neurotransmitters (Kolodziejczyk et al., 2008; Mauss et al., 2014) that perform signaling in the fly’s visual system are acetylcholine, GABA, and glutamate. Unlike in the mammalian brain, where glutamate functions as an excitatory neurotransmitter, glutamate can also be inhibitory in the fly nervous system, acting via a glutamate-gated chloride conductance (Liu and Wilson, 2013).

1.4 Tools to Study *Drosophila* Vision

Merits of *Drosophila* as a model organism include the variety of methods that have been introduced and can be applied in the research, as having a large toolbox to dissect the anatomy and function of the visual circuit is indispensable.

Below, I describe some of the most commonly used methods, with a specific focus on the ones that were central to my PhD research and to the publications presented in this thesis.

1.4.1 Targeting Cells: The Gal4/UAS system

One of the main tools in the *Drosophila* toolbox is the binary Gal4/UAS (Figure 3) expression system that allows genetic access to the target cell types and single neurons (Brand and Perrimon, 1993).

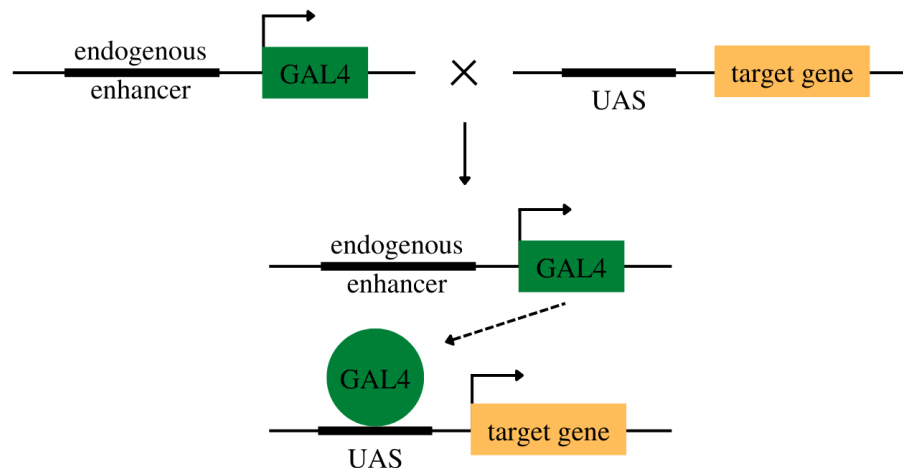


Figure 3. Gal4/UAS System.

Gal4/UAS system consists of two parts: one indicating “what” should be expressed and the other “where” it should be expressed. Or, respectively, Upstream Activating Sequence (UAS), which is an enhancer sequence in the DNA, and Gal4, a yeast transcription factor that it binds to. Strains of flies with only Gal4 are called the “driver” lines and do not produce any

effects by themselves as *Drosophila* has no endogenous UAS. However, combining the “driver” line with a fly strain that has UAS next to the target region (the “reporter” line) creates a powerful tool that allows targeting the expression of a specific reporter solely to the neurons of interest.

With this tool, libraries of thousands of “reporter” lines have been established (for example, Jennet et al., 2012), providing a variety of unique cell-specific expression patterns.

1.4.2 Visualizing Cells and Cell Activity: GFP and GCaMP

The “reporter” genes expressed through the Gal4/UAS system can be used to express fluorescent proteins, which are excited with light of specific wavelengths, and visualize the structure, as well as the activity of the cells of interest (see Figure 4 for example images).

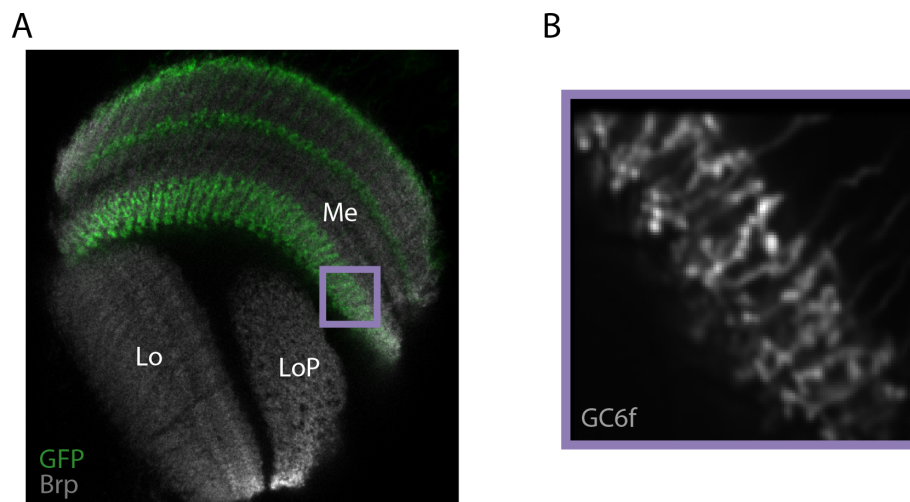


Figure 4. Visualizing Cells and Cell Activity. A. Mi1 GFP staining. B. GCaMP6 signal in Mi1 cells in layer M10.

An example of the former is the Green Fluorescent Protein, GFP (Chalfie et al., 1994). It is derived from a jellyfish and, when expressed in the cell of interest, visualizes its structure, allowing, for example, to localize the soma and cell's arborizations.

To visualize the activity of the cell, one proxy that is often used is the fluorescence of indicators that show the presence of different chemical elements and neurotransmitters in the cell. Indicators have even been developed to directly access the cell's voltage: so-called GEVIs, genetically encoded voltage indicators (Cohen, 1989). However, one of the most commonly used indicators up to date remains the genetically encoded fluorescent sensor for Ca^{2+} , which measures the intracellular concentration of calcium as a proxy for neuronal

activity. Initially developed in 2001 as GCaMP1 (Nakai et al., 2001), it has undergone many improvements to become the commonly used GCaMP6 (Chen et al., 2013) and the most recently created GCaMP8 (Looger et al., 2021), and continues to develop further, established as an irreplaceable tool in the neuron activity visualization in a variety of model organisms. When it is expressed in a specific cell type, for example, via the Gal4/UAS system, it can be used to optically monitor the activity of a neuron or, even more specifically, a neuronal compartment, such as the cell body or the dendrites.

However, calcium indicators also possess a number of drawbacks. Firstly, they are inherently slow, serving as low pass filters for the neuronal activity and making the reported signal slower in comparison to the membrane voltage changes, as they have their own rates of binding and release of calcium, i.e., dissociation constants, which limit their temporal response dynamics. This, in turn, affects the concentration of the free calcium available to the cell, influences calcium dynamics in the cases of high calcium concentration, and can cause nonlinear perturbations (Borst and Abarbanel, 2007). Secondly, intracellular calcium concentration is an indirect indicator of neuronal activity and does not one-to-one translate into membrane voltage: it cannot detect sub-threshold depolarization, rectifies the signal, and can only report the changes in the membrane voltage that lead to significant calcium changes (Chamberland et al., 2017).

1.4.3 Manipulating the Circuit: TeTxLC

There is, however, no need to remain passive observers: the available genetic toolbox also makes it possible to manipulate the visual circuit. This can be achieved by, for example, interfering with the synaptic transmission mechanisms, affecting the release of the neurotransmitters, and thereby silencing the target cells. One of the tools that can be used to block synaptic transmission is Tetanus Toxin Light Chain (TeTxLC), which cleaves the synaptic vesicle protein synaptobrevin, disrupting the synaptic vesicle release (Sweeney et al., 1995).

Manipulating the circuit in such a way helps answer questions that would not be interpretable through observation of the cells' activity alone. For example, it allows for testing the extent to which the response of a cell depends on the feedback coming from within the same cell type, even when the signal needs to go through an extra cell in between before feeding back into the original cell, or imaging downstream cells and dissecting the importance of individual cell types as inputs into the circuit.

1.4.4 Recording Single Neuron Activity: Electrophysiology and Two-Photon Microscopy

Two methods most commonly used to record neural activity in *Drosophila* are whole-cell patch-clamp recording and two-photon excitation microscopy (Figure 5A and 5B, respectively), both often employed to understand electrical signaling and signal propagation between neurons.

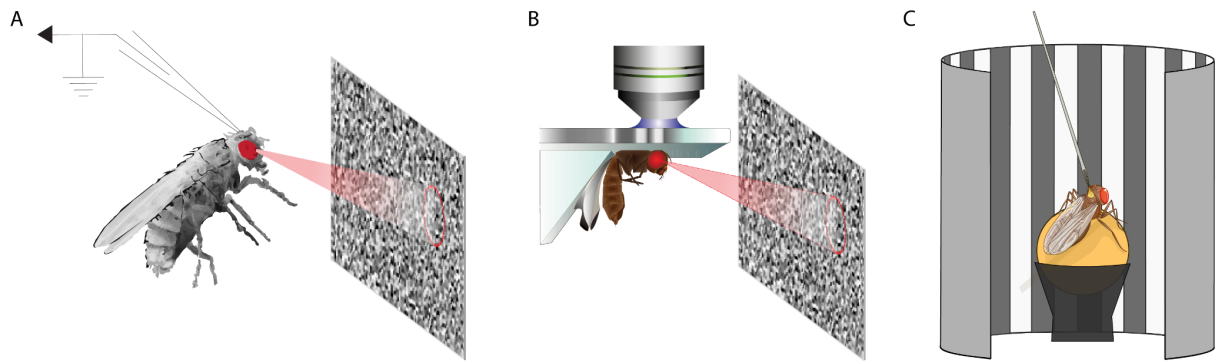


Figure 5. Tools to Study *Drosophila* Vision. Schematics of setups for: A. Electrophysiology, B. Two-photon microscopy, and C. Tethered walking.

Whole-cell patch-clamp recordings (Sakmann and Neher, 1984) allow direct measurement of voltage fluctuations from a single cell body in living tissue. It is the most direct way of accessing neural activity, unhindered by indicator dynamics. While the success of whole-cell patch-clamp recordings is limited by the size of the soma, they have been used successfully in different brain regions from fixed, walking, and flying flies (for example, olfactory responses in fixed flies were measured by Wilson et al. (2004), and visual responses during tethered flight by Maimon et al. (2010)). Recently, it has also become feasible to carry out patch-clamp recordings from small, previously inaccessible neurons with a soma size of less than 3 μm (Gruntman et al., 2018; Groschner et al., 2022)

Although electrophysiology offers the most direct access to cell signaling, it comes with certain limitations: recording from the cell body might not necessarily reflect the neural signal in other neuronal compartments in the cases of larger, highly compartmentalized neurons. Additionally, even though access to neurons of smaller size is now possible, anatomical accessibility of neurons still imposes certain limitations, such as the difficulty of accessing cells with their somas located under another dense cellular layer.

Two-photon excitation microscopy is an imaging technique that allows optical access *in vivo* to fluorescent indicators expressed in the cells of interest (Denk et al., 1990). Here, the

simultaneous absorption of two low-energy photons of near-infrared light by a fluorophore excites it so that it then emits light in the visible range.

Two-photon excitation microscopy allows access to cells that might be inaccessible to electrophysiology because of their size or position. It provides good temporal resolution of the signal (on par with the resolution of the indicators used), enough spatial resolution to distinguish individual neurons and their compartments, access to living tissue, and, in addition to the single-cell access offered by electrophysiology, it also allows simultaneous recordings from a large number of neurons.

1.4.5 Observing Natural Behavior: Behavioral Assays

Fruit flies exhibit an optomotor response: a consistent pattern of behavior, whereby when the environment around the fly is turning in a certain direction, the fly will also turn in that direction. This way, the fly successfully compensates for, what is perceived as, being blown off course by the wind. The magnitude of this turning response is determined by the internal estimate of velocity, and this estimate has been shown to be specifically tailored for natural environments (Chen et al., 2019). The optomotor response serves course stabilization and can be exploited in behavioral experiments, such as when testing flies on tethered flight and tethered walking (ball) setups (Götz, 1964) (Figure 5C).

In the walking setup, a tethered fly is placed onto a ball, which lies on an air cushion and thus can freely turn with the movement of the fly. The fly is surrounded by screens that can deliver visual stimuli, intended to induce an optomotor response: for example, if a grating moving in a certain direction is displayed, the fly would start walking in the same direction, moving the ball it is walking on. The displacement of the ball can be measured and used as a readout of the fly's behavior, corresponding to the yaw rotation of the fly. Comparing the direction and speed of the ball rotation with the direction and speed of the stimulus rotation makes it possible to draw conclusions about the fly's internal estimation of the motion.

Tethered walking setups provide numerous advantages, such as providing access to the natural behavior of the fly and, thanks to their high throughput, making it possible to carry out fast screenings for mutations, characteristics, or stimuli that might be of interest. The main drawback of the tethered walking experiments is the difficulty of unambiguous translation of the results obtained from the ball setups into flying behavior.

1.4.6 Recreating the Circuit: Modeling

Knowing the individual elements and their properties, one might want to reconstruct the circuit analytically. Matching the already known characteristics to the theoretical models makes it possible to formulate hypotheses about functional architectures, find missing components of a circuit, and instruct further experimental approaches.

Modeling can be performed at different levels from true to the biological reality, biophysical and compartmentalized models, to a higher level, conceptual or algorithmic circuit models with simplified elements.

Models play an important role in circuits neuroscience and help us find a correspondence between computational modules and neural mechanisms in the brain.

1.5 Studying Individual Elements of the Motion Detection Circuit

1.5.1 Computations in the *Drosophila* Optic Lobe

As the signal is transmitted through the neuropils down the optic lobe, computations performed by the cells increase in their complexity.

One way to categorize such computations is by calculating the cell's linear filter, i.e., by extracting its linear spatiotemporal receptive field (RF) (Borst et al., 2020). This can be done by reverse-correlating a neuron's response to a stochastic stimulus, such as pixel or bar noise (Eggermont et al., 1983; Chichilnisky, 2001). While the result is a linear approximation that discounts any potential non-linearities in the cell's computations, it is nevertheless a reliable representation of the filtering operations that the cell performs. Using the spatiotemporal RF of the cell, one can determine the cell's preference for the stimulus polarity, find the spatial location of its receptive field on the stimulus screen, and categorize the cell's impulse response, i.e., measure its immediate response to a brief input signal. This technique is often used to categorize visual neurons in the *Drosophila*'s optic lobe and specifically within the motion vision circuit (see Arenz et al., 2017 for linear spatiotemporal RFs of the main inputs into the motion vision circuit).

In the retina, a photoreceptor responds merely to the number of photons it receives and processes the incoming signals by scaling them to fit the limited output range of the downstream neurons.

In the lamina, non-spiking neurons respond either transiently or sustained to both light and dark edges (specifically, responding with hyperpolarization to luminance increases and depolarization to luminance decreases), and are not selective for moving stimuli (Tuthill et al., 2013; Borst et al., 2020). As the output to the downstream neurons, the lamina transmits information about the relative luminance difference, i.e., the absolute local contrast.

The separation of signal processing into ON and OFF pathways arises in the lamina (Joesch et al., 2010; Maisak et al., 2013), with cells within a pathway responding exclusively either to increments or to decrements in luminance. This division continues robustly in the medulla and further along the motion vision circuit (with the exception of Mi9 cells that respond to luminance decrements despite being a part of the ON visual pathway). Medulla cells also differ from each other with respect to the size and complexity of their RFs: at this stage, cells have larger RFs, many of them with prominent antagonistic center-surround components (Arenz et al., 2017). Here, the division of the cells into low- and band-pass filters with a spectrum of time constants becomes apparent, and the transient and sustained characteristics of the responses produced by these cells play an important role in the motion detection computations in the downstream circuit. The medulla corresponds to the second processing stage, where information about the absolute local contrast, received from the lamina, is processed to represent the relative contrast and transmitted further downstream.

The medulla and the lobula also house the dendrites of the first direction-selective neurons. These cells are even more selective about the stimuli they respond to and now extract higher-order visual cues from their inputs.

Within the lobula plate, neurons discriminate visual features further. Here, the lobula plate tangential cells (LPTCs) integrate over large parts of the visual field and respond with depolarization to visual motion in their preferred direction and hyperpolarization to motion in the opposite direction (Hausen, 1984). From the lobula plate, the information from the motion-sensitive pathways is projected further to the protocerebrum.

Going down the processing pathway in the optic lobe of *Drosophila*, from the retina to the lobula plate, one can see the increase in the complexity of the computations that the cells perform, as well as in the rise in the complexity of the information that is transmitted at the output of each stage.

1.5.2 Motion Detection Circuit in *Drosophila* Optic Lobe

Detecting motion and its direction are fundamental tasks, however, the direction in which the image is shifting is not explicitly represented at the level of individual photoreceptors, as

a single photoreceptor only contains information about luminance change at one location in space. Rather, directional motion information needs to be extracted from the photoreceptor array by comparing the signals of neighboring units.

In the fly optic lobe (Figure 2) none of the medulla cells are directionally selective, and the dendrites of T4 and T5 cells (T4/5), so-called bushy T cells, are the first stage to represent the direction of motion (Maisak et al., 2013; Fisher et al., 2015). The two cell types process motion information in parallel and independently, with T4 being a part of the ON pathway and T5 its counterpart in the OFF pathway (Joesch et al., 2010; Maisak et al., 2013; Fisher et al., 2015). T4 cells have their dendrites in the medulla layer 10, T5 cells in the posterior layer Lo1 of the lobula, and both T4 and T5 cells send their projections to the four layers of the lobula plate.

The ON and OFF motion pathways are similar in their function, characteristics of their main inputs, and patterns of synaptic connections. Both T4 and T5 cells are further divided into four subtypes: a, b, c, and d, each sensitive to motion in one of the four cardinal directions (Maisak et al., 2013). Each subtype projects its axon to one of the lobula plate's four layers, depending on the direction of motion that it signals: the subtype that responds to front-to-back motion projects to layer 1, back-to-front to layer 2, and upwards and downwards responding subtypes project to layers 3 and 4, respectively (Maisak et al., 2013).

Takemura et al. (2017) comprehensively revealed medulla neurons that provide inputs to T4. Shinomiya et al., (2014) also identified the first candidates for the OFF motion detection pathway. All inputs, their relative proportions, and the location where they contact T4/5 dendrite have been identified (Shinomiya et al., 2019), with the main T4 inputs being, in descending order starting from the ones with the most synaptic contacts: Mi1, Tm3, Mi9, CT1, Mi4, and TmY15. Main inputs to T5 are as follows: Tm9, Tm2, Tm1, CT1, Tm4. Additionally, T4/5 a-d subtypes also show recurrent connections within a subtype. Combining this dense EM reconstruction (Takemura et al., 2017) with the later characterization of the linear spatiotemporal RFs of the input neurons (Arenz et al., 2017) also revealed that T4 and T5 cells receive input from tonic neurons (such as Mi4, Mi9, and Tm9) that carry a slow, temporally low-pass filtered signal within one column, as well as from transient neurons (such as Mi1, Tm1, Tm2, Tm3, and Tm4) that carry a fast, temporally band-pass filtered signal within adjacent columns (Borst et al., 2020).

1.5.3 Coincidence Detectors

As individual photoreceptors cannot detect motion, motion detection requires information from at least two neighboring photoreceptors, which send their outputs to a coincidence detector that can further read out the direction of motion from those signals. The main rules when building such a coincidence detector (Borst et al., 2020) are: 1. It needs to receive at least two inputs that are offset in space; 2. These inputs should be asymmetrically filtered in time to allow the coincidence to happen; 3. The signals arriving from the two inputs have to interact with each other in a non-linear way, allowing the direction of motion to be extracted (specifically, there should be a response to the motion in one direction and no response to the motion in the other). To complete the full elementary motion detector (EMD), two such coincidence detector units that are mirror-symmetrical to each other are required (Reichardt 1987; Borst and Egelhaaf 1989). Subtracting the signal of two subunits with the opposite direction preference produces a fully opponent response: consequently, producing a positive response to motion in one direction and a negative one in response to motion in the other direction.

In the case of the fly visual system, the cells' dendritic arbors receive parallel inputs from multiple columns, with a single arbor receiving inputs from columns signaling different positions in the visual field, depending on the cell types of the input neurons (Takemura et al., 2017). These signals are non-linearly combined within T4 and T5 cells so that stimuli traveling along one direction (the so-called 'preferred direction') are amplified, and stimuli in the opposite direction (the so-called 'null direction') become suppressed.

What computation happens between these signals determines the nature and the directional preference of the detector (Figure 6). Historically, two of the most famous detectors that have been used to explain motion detection in *Drosophila* are Hassenstein–Reichardt (HR; Hassenstein and Reichardt, 1956) and Barlow–Levick (BL; Barlow and Levick, 1965) detectors. Recently, a new detector has been added to the collection, created as a combination of the two: Hassenstein–Reichardt/Barlow–Levick (HR/BL; Arenz et al., 2017) detector.

Within each subunit of the HR detector, direction selectivity is achieved through "preferred direction enhancement". This means that when the two input signals (delayed and non-delayed) arrive at the non-linear processing stage simultaneously, which happens when the delayed signal gets activated first, the signals enhance each other via an operation equivalent to multiplication. The corresponding direction of motion is, therefore, the

detector's preferred direction. In other words, the HR detector enhances responses to motion in the preferred direction by combining offset excitatory inputs.

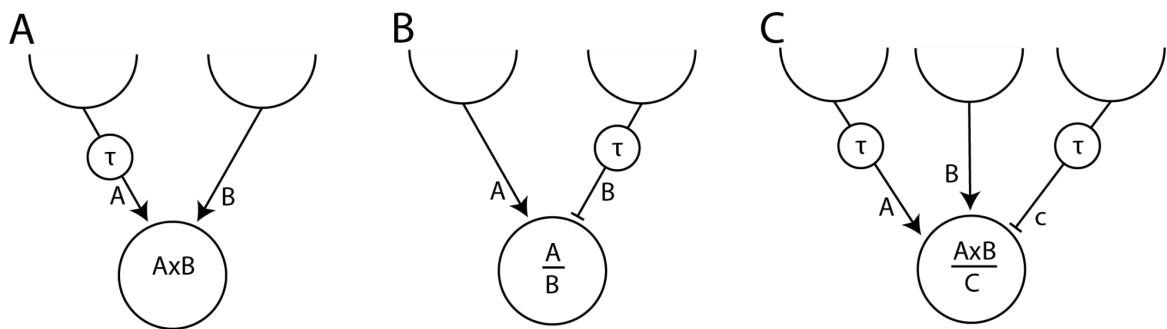


Figure 6. Motion detectors. Examples of A. Hassenstein-Reichardt, B. Barlow-Levick, and C. Hassenstein-Reichardt/Barlow-Levick motion detectors.

In the BL detector, direction selectivity is achieved through “null direction suppression”. Here, activity is suppressed when both input signals arrive at the non-linear processing stage simultaneously. This operation is equivalent to a division, and the corresponding direction of motion is, therefore, the detector's null direction. Or, biologically, the BL detector uses inhibitory input to “veto” an offset excitatory input, suppressing motion in the null direction (Gruntman et al., 2018).

Recently proposed HR/BL detector combines both HR and BL detectors and, consequently, accounts rather well for both preferred direction enhancement on the preferred side and null direction suppression on the null side of the dendrite, thus reflecting the basic response properties of T4 and T5 cells. In particular and in contrast to either detector alone, the model captures the high degree of direction selectivity that is observed in T4 and T5 cells right at the first stage, where direction selectivity emerges (Haag et al., 2016; Leong et al., 2016; Arenz et al., 2017). Intriguingly, this algorithmic input structure is rather well reflected by the anatomic arrangement of presynaptic cells and their temporal filtering properties: fast Mi1 and Tm3 are positioned in the central part of the T4 dendrite, while slow neurons Mi4/CT1 and Mi9 are located at the null and preferred flanking sides, respectively. For T5, similarly, fast Tm1, Tm2, and Tm4 are centrally located, while slow CT1 and Tm9 act on the null and preferred sides.

The core process of direction selectivity as implemented on the dendrites of T4 and T5 cells comprises both an enhancement of signals for motion along their preferred direction and a suppression of signals for motion along the opposite direction (Haag et al., 2017). This

combined strategy ensures a high degree of direction selectivity right at the first stage where the direction of motion is computed.

1.5.4 Beyond the First Stages of Motion Detection

At the subsequent motion detection processing stage, information about behaviorally relevant optic flow fields is extracted by a set of wide-field neurons called lobula plate tangential cells (LPTCs). LPTCs' characteristic feature is their motion opponency: they respond to motion along their preferred direction with depolarization and they hyperpolarize when stimulated by motion in the opposite direction (Hausen 1976; Hausen, 1984; Joesch et al., 2008), just like the subtraction stage of the fully opponent motion detector.

LPTCs derive their flow-field sensitivity from integrating direct excitatory, cholinergic signals from ON and OFF selective T4 and T5 cells and indirect inhibition from bi-stratified lobula plate intrinsic neurons (LPi) activated by neighboring T4/T5 terminals. These signals are integrated onto the large LPTCs dendrites within selected layers of the lobula plate (Schnell et al., 2012; Mauss et al., 2014), and, as a result, LPTCs are more specifically tuned to particular flow fields and do not respond to other flow fields, even when they partially match their preferred flow field within certain patches of the visual field (e.g., translation versus expansion) (Mauss et al., 2015).

1.6 Scopes and Aims of the Thesis

Contrast normalization is a canonical computation found in the visual systems of various vertebrates, that keeps the neural representation of stimuli robust despite varying contrast conditions in the environment. The role of contrast normalization in the early visual system of *Drosophila*, however, and specifically within the motion detection circuit and its main inputs, has not yet been described.

During my PhD, I studied the effects of contrast normalization on neuronal processing and motion detection, as well as its effect on the filtering properties of the neurons.

In my first project, as described in Drews et al. (2020), we focused on the question: “What role does contrast normalization play in the motion detection circuit of *Drosophila*?”. We found at which stage of the early visual processing in the optic lobe contrast normalization first arises, showed the suppressive effect that it has on the amplitude of the signal, and categorized the major inputs of the motion detection circuit into those affected by contrast normalization and the ones that are not. We suggested possible mechanisms via which

contrast normalization is implemented and showed the importance of neural feedback. We also studied the effect that contrast normalization has on motion detection and concluded that it improves the performance of the motion detection circuit in response to natural stimuli. Additionally, we found that, when included in a motion detection model that would normally underperform when compared to the responses of single cells, contrast normalization improves the model's performance. It allowed us to fine-tune the model, and might, thus, be the missing element in our understanding of motion detection.

In my second project, described in Pirogova and Borst (in revision), we aimed to answer the question: “What effect does contrast normalization have on the dynamics of the neurons in the motion detection circuit of *Drosophila*?”. We investigated the impact of static and dynamic surround stimuli on the cell’s responses to a central stimulus and determined that it affected the temporal dynamics of the signal, in addition to suppressing the signal’s amplitude. We asked how the effects of contrast normalization influence the filtering properties of the visual neurons when they are categorized with artificial stimuli. We also further narrowed down the search for the mechanism behind contrast normalization in the motion detection circuit of *Drosophila*.

Overall, we demonstrated the presence of contrast normalization early in the *Drosophila* optic lobe, showed that it suppresses response amplitude and increases response speed, confirmed that it partially relies on the neuronal feedback, narrowed down the search for the implementation mechanism, and demonstrated that normalization improves motion detection in the model responses to natural stimuli.

Current Biology

Dynamic Signal Compression for Robust Motion Vision in Flies

Highlights

- *Drosophila* motion processing robustly estimates the velocity of moving natural scenes
- Visual interneurons in the fly visual system dynamically adapt to stimulus contrast
- This adaptation relies on fast spatial integration of neural feedback
- Contrast adaptation accounts for robust motion vision in computational circuit models

Authors

Michael S. Drews, Aljoscha Leonhardt, Nadezhda Pirogova, ..., Lukas Braun, Etienne Serbe, Alexander Borst

Correspondence

drews@neuro.mpg.de (M.S.D.),
leonhardt@neuro.mpg.de (A.L.)

In Brief

Flies reliably estimate the velocity of moving natural scenes regardless of image statistics. Current models of *Drosophila* motion vision fail to explain this robustness. Drews, Leonhardt, et al. show that flies achieve this performance by rapidly adjusting the sensitivity of visual interneurons in the medulla to surround contrast.



Dynamic Signal Compression for Robust Motion Vision in Flies

Michael S. Drews,^{1,2,4,*} Aljoscha Leonhardt,^{1,4,5,*} Nadezhda Pirogova,^{1,2} Florian G. Richter,^{1,2} Anna Schuetzenberger,^{1,2} Lukas Braun,³ Etienne Serbe,¹ and Alexander Borst¹

¹Department Circuits-Computation-Models, Max-Planck-Institute of Neurobiology, 82152 Martinsried, Germany

²Graduate School of Systemic Neurosciences, LMU Munich, 82152 Martinsried, Germany

³Bernstein Center for Computational Neuroscience, 10115 Berlin, Germany

⁴These authors contributed equally

⁵Lead Contact

*Correspondence: drews@neuro.mpg.de (M.S.D.), leonhardt@neuro.mpg.de (A.L.)

<https://doi.org/10.1016/j.cub.2019.10.035>

SUMMARY

Sensory systems need to reliably extract information from highly variable natural signals. Flies, for instance, use optic flow to guide their course and are remarkably adept at estimating image velocity regardless of image statistics. Current circuit models, however, cannot account for this robustness. Here, we demonstrate that the *Drosophila* visual system reduces input variability by rapidly adjusting its sensitivity to local contrast conditions. We exhaustively map functional properties of neurons in the motion detection circuit and find that local responses are compressed by surround contrast. The compressive signal is fast, integrates spatially, and derives from neural feedback. Training convolutional neural networks on estimating the velocity of natural stimuli shows that this dynamic signal compression can close the performance gap between model and organism. Overall, our work represents a comprehensive mechanistic account of how neural systems attain the robustness to carry out survival-critical tasks in challenging real-world environments.

INTRODUCTION

Visual motion represents a critical source of sensory feedback for navigation. Self-motion results in particular patterns of local directional cues across the retina. Detection of these optic flow fields allows animals to estimate and control their current heading [1]. Flies, for instance, react to whole-field retinal motion by turning in the same direction as their surroundings. This optomotor response enables them to maintain a straight path under perturbations as well as over long distances [2, 3].

For the reflex to work effectively, biological motion detectors need to respond reliably and independently of the particular visual statistics of the environment. This poses a challenge given the complexity of natural scenes [4, 5]. Motion vision systems therefore need to employ processing strategies that maintain robust performance despite the variability of natural visual input.

Recent circuit mapping efforts have yielded unprecedented insight into the neural substrate of motion detection in *Drosophila* [6, 7]. The fly optic lobe consists of sequential neuropils (retina, lamina, medulla, lobula, and lobula plate) and is arranged in columns that process visual input retinotopically. In various combinations, lamina cells L1–L5 feed into a light-sensitive ON or a dark-sensitive OFF pathway, each comprising at least four cell types in the medulla [8]. Medulla units fall into two classes characterized either by transient temporal filtering and moderate center-surround antagonism in their spatial receptive field (Mi1 and Tm3 for ON; Tm1, Tm2, and Tm4 for OFF) or by tonic responses and strong antagonistic surround (Mi4 and Mi9 for ON; Tm9 for OFF) [9–13]. Postsynaptic T4 and T5 cells then compute local ON and OFF motion, respectively, by comparing medulla signals with different dynamics across neighboring columns [8, 14–19]. Jointly, they are necessary for the optomotor response [20]. By pooling appropriate T4 and T5 signals, lobula plate tangential cells (LPTCs) detect optic flow fields that correspond to rotations around different body axes and ultimately control turning [3, 21–23].

For artificial stimuli, fly motion processing is well explained by correlation-based detector models that rely on multiplication of spatially adjacent, asymmetrically filtered luminance signals [24]. These elementary motion detectors (EMDs) account for subtle features of behavioral and neural responses such as pattern-induced shifts in velocity tuning [25, 26], intrinsic velocity gain control [27], or reverse-phi sensitivity [28, 29]. However, EMD output strongly depends on contrast as defined by the average difference between light and dark [26]. EMDs thus invariably confound image contrast with velocity. Since local contrast varies substantially within natural images [4], output from individual EMDs is sparse and fluctuates heavily under naturalistic conditions (Figures S1A–S1C). Motion responses in flies, however, have been shown to be highly robust, across both time and different natural scenes [30, 31].

Various general mechanisms for adaptation to naturalistic signals have been described in the fly visual system. These include gain control in photoreceptors or LPTCs [32–34], redundancy reduction through lateral inhibition [35], subtractive enhancement of flow field selectivity [36], and tailoring of processing to fundamental natural scene statistics [31, 37, 38]. However, none effectively address the problem of contrast fluctuations.



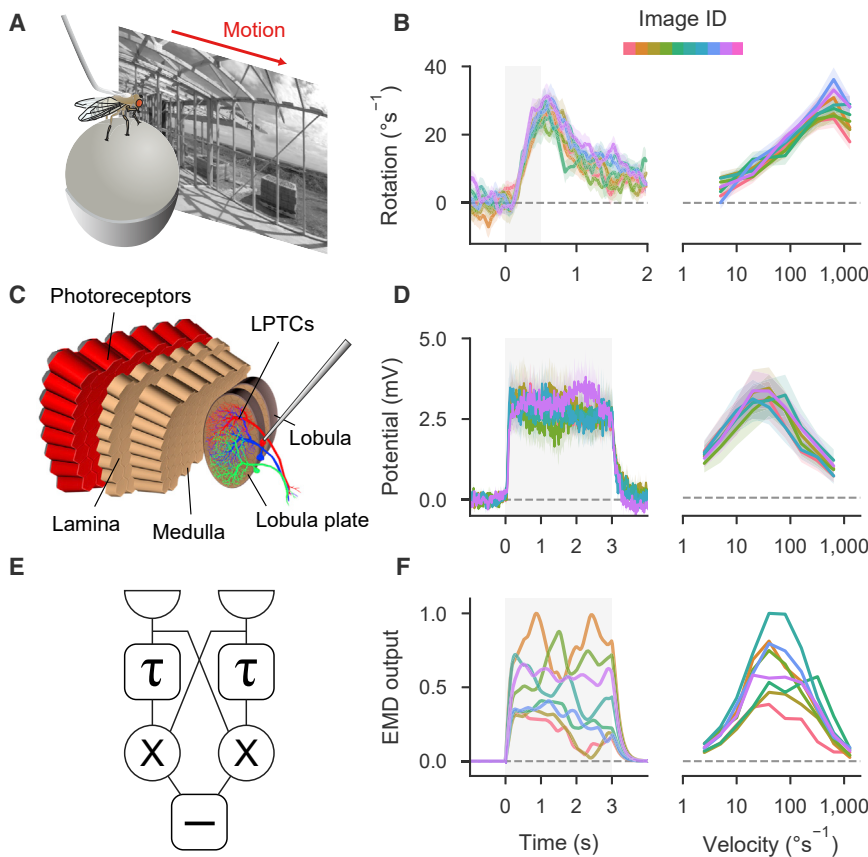


Figure 1. Flies Respond More Robustly to Natural Scene Variability Than Predicted by Correlation-Based Motion Detectors

(A) Illustration of behavioral set-up. Tethered wild-type *Drosophila* were stimulated with translating natural images.

(B) Left: turning responses for images moving at 80°s^{-1} ($n = 16$ flies). Each color indicates a distinct scene. Images moved during gray-shaded period. Right: velocity tuning curves for all measured scenes (averaged between 0 and 1 s after motion onset).

(C) Illustration of fly visual system. Photoreceptor signals are processed in five retinotopically arranged neuropils. Wide-field lobula plate tangential cells (LPTCs) respond to particular optic flow fields.

(D) Left: membrane potential of horizontal system LPTCs in response to images moving at 20°s^{-1} ($n = 11$ cells from 9 flies). Right: velocity tuning curves (averaged between 0 and 3 s after motion onset).

(E) Schematic of an individual correlation-based elementary motion detector (EMD; τ denotes delay line; \times , multiplication; $-$, subtraction).

(F) Left: responses of an array of EMDs to stimulation with natural images moving at 20°s^{-1} . Right: velocity tuning curves of EMD array (evaluated like LPTC output). Note that in contrast to experiments, model responses were averaged across many different starting phases. Shaded areas around curves indicate bootstrapped 68% confidence intervals.

See also Figure S1 and Table S2.

In vertebrate visual systems, contrast sensitivity is continuously regulated through the mechanism of divisive normalization [39–41]. Here, the response of a neuron is effectively divided by local contrast, estimated as the average activity within a population of neighboring neurons. The process compresses signals of varying contrast into a fixed range by dynamically adjusting gain to current conditions [5] and renders the neural representation of stimuli largely invariant with respect to contrast. However, so far, no comparable mechanism has been described for the invertebrate visual system.

Here, we investigate how the fly visual system copes with contrast variability and demonstrate that dynamic signal compression based on divisive contrast normalization renders motion processing robust to the challenges imposed by natural visual environments.

RESULTS

Fly Motion Responses Are Robust to Natural Scene Variability

To rigorously assess the robustness of *Drosophila* motion processing, we measured optomotor responses to a diverse set of moving naturalistic panoramas on a walking treadmill setup (Figure 1A). Fly turning was highly consistent across images and velocity tuning curves showed virtually no variation over different scenes, matching previous findings [31] (Figure 1B; Figure S1). To quantify reliability at the neural level, we recorded the membrane potential of horizontal system LPTCs that detect

optic flow fields corresponding to yaw rotation (Figure 1C). Potential was tuned to scene velocity and again exhibited little image-dependent variation (Figure 1D). Additionally, membrane voltage proved highly stable across time. This was consistent with earlier work in hoverflies [30].

To perform a consistent comparison, we tested the robustness of EMDs on the same set of stimuli as in behavior and electrophysiology (Figure 1E). As anticipated from a multitude of similar studies [31, 37, 42, 43], responses were remarkably unreliable across time and images (Figure 1F). For most images, temporally resolved output fluctuated strongly, average amplitudes differed, and tuning curves exhibited peaks at different velocities. Overall, EMDs provided a poor readout of true image velocity. This stands in stark contrast to the experimentally observed robustness of motion responses and leads to the central question: how does the fly visual system compensate for natural contrast variability?

Sensitivity of Optomotor Response Is Modulated by Surround Contrast

We designed an optomotor stimulus to establish whether *Drosophila* dynamically adapt the sensitivity of motion-induced turning to image contrast, which could serve to normalize variation within natural scenes. The stimulus segregated the visual field into a background and a foreground pattern (Figure 2A). The background contained random luminance fluctuations but no net motion. Pattern movement within the foreground window triggered turning. For both, average contrast could be controlled

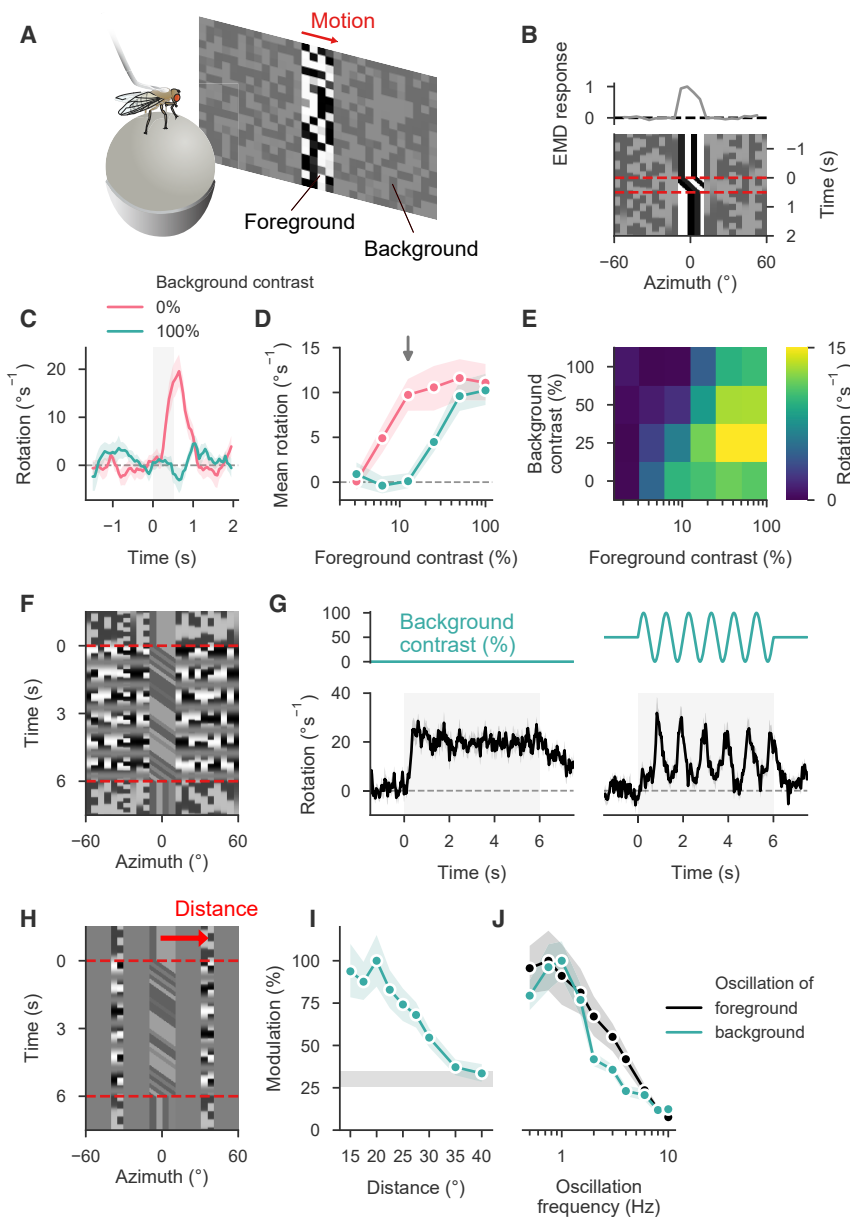


Figure 2. Sensitivity of *Drosophila* Optomotor Response Is Controlled by Surround Contrast

(A) Experimental set-up. Visual display is separated into two areas whose contrast can be set independently.

(B) Bottom: space-time plot of base stimulus. Foreground pattern moved during time span indicated by dashed lines; background is dynamic but contains no coherent motion. Top: time-averaged response of EMD array along azimuth. Only foreground produced net activity.

(C) Turning responses for extreme background contrast conditions ($n = 16$ wild-type flies) at foreground contrast 12.5%. Gray-shaded area indicates motion.

(D) Mean rotation (averaged between 0 and 1 s after stimulus onset) as a function of foreground contrast for two background conditions ($n = 16$; gray arrow indicates foreground contrast depicted in C).

(E) Heatmap of mean rotation for multiple background conditions. With increasing background contrast, optomotor sensitivity shifted rightward ($n = 16$).

(F) Example stimulus for mapping magnitude of sensitivity shift. Background contrast was modulated at 1 Hz.

(G) Left: baseline turning response in the absence of background contrast ($n = 16$, foreground contrast 25%). Right: turning response for sinusoidal change in background contrast (data taken from spatial experiment evaluated in I at distance 15°). During high-contrast phase, optomotor response was suppressed; turning modulation allowed readout of background-induced changes in gain.

(H) Illustration of spatial oscillation experiment. Distance indicates separation between centers of foreground motion and flanking background.

(I) Turning response modulation as a function of distance between motion stimulus and background ($n = 16$). Gray-shaded bar indicates 68% confidence interval around baseline modulation in the absence of background.

(J) Turning response modulation as a function of carrier frequency for either foreground ($n = 13$) or background ($n = 13$). Shaded area around curves indicates bootstrapped 68% confidence interval. See also [Figure S2](#), [Table S2](#), and [Videos S1](#) and [S2](#).

independently. We confirmed that the background by itself produced no net activity in EMDs (Figure 2B).

At zero background contrast, foreground motion induced a reliable optomotor response (Figure 2C). Turning was fully suppressed at maximum background contrast, proving that turning gain is controlled by surround contrast. Average field luminance was constant for all conditions, so linear processing could not account for the phenomenon. A full measurement of contrast tuning curves for foreground motion revealed a smooth shift of the dynamic range of the optomotor response toward the current surround contrast (Figures 2D and 2E).

To efficiently map features of contrast gain control in a single stimulus condition, we sinusoidally modulated background contrast over time, which resulted in oscillations around mean turning (Figures 2F and 2G). Whenever background contrast was high, syndirectional rotation in response to motion was

transiently suppressed. Evaluating oscillation amplitude thus allowed a readout of the level of contrast-induced gain adjustment. We determined the spatial scale of suppression by varying the spacing between foreground and a windowed background, separated by uniform gray (Figure 2H). Modulation fell with distance between motion stimulus and background stripe and dropped to baseline at approximately 35° , so contrast estimation was non-local but spatially limited (Figure 2I; Figures S2A and S2B; full width at half maximum of 43.8° for zero-centered Gaussian least-squares fit to mean tuning curve).

When we varied oscillation frequency in the background, suppression followed contrast changes up to fast timescales beyond 3 Hz (Figure 2J; Figures S2C–S2F). However, modulation decreased at lower frequencies than for equivalent foreground oscillations, which is indicative of temporal integration. We additionally evaluated the lag between contrast oscillation and

turning by means of cross-correlation (Figures S2G and S2H). The maximum suppressive effect of background modulation was delayed with respect to the effect of foreground modulation by approximately 70 ms (bootstrapped 95% CI: 33–114 ms). This supported the previous conclusion that the mechanism for background contrast estimation operates on slower timescales than the primary motion pathway. Silencing T4 and T5 cells abolished all contrast-guided oscillatory turning (Figures S2I–S2K), suggesting that contrast adaptation is not mediated by a system parallel to motion detection [44]. Our experiments thus point to a rapid, spatially distributed gain control mechanism that arises in early visual processing.

Signal Compression Emerges in Transient Medulla Neurons

We next used two-photon calcium imaging to locate the neural origin of contrast adaptation. The calcium indicator GCaMP6f was genetically expressed in particular cell types [45]. We targeted visual stimuli to individual neurons by determining receptive field coordinates through a combination of stochastic stimuli and online reverse correlation (Figures 3A and 3B; STAR Methods). This procedure additionally yielded estimated linear receptive fields for L1–L5, analogously to the ones previously described for medulla neurons [9] (Figures S3A–S3T). Consistent with earlier functional work [29, 46], spatiotemporal filters grouped into tonic (L3) or transient units (L1, L2, L4, and L5) like they did in the medulla. In contrast to all other lamina cells, we found that the polarity of the L5 receptive field center is ON.

To precisely map context-dependent changes in contrast sensitivity for a given cell type, we then presented drifting sine gratings with separately controlled contrast in the foreground (as defined by a 25° circular window centered on the receptive field) and the background (Figure 3C). At a fixed foreground contrast, L1 activity followed local grating luminance and was independent of background contrast (Figure 3D). Responses in downstream synaptic partner Tm3, however, showed the signature of gain control as signal amplitude was increasingly suppressed by growing surround contrast (Figure 3E).

We performed these experiments for all major columnar cell types in the circuit as well as T4 and T5 cells (Figure 3F). To obtain contrast tuning curves, we evaluated calcium modulation at the stimulus frequency. Lamina units tracked foreground contrast but were weakly, if at all, modulated by the surround except for a vertical shift at low levels (Figures 3G–3K). This was likely due to background leaking into the receptive fields since antagonistic surrounds extend beyond 25° for some cell types (Figure S3) [9]. In the medulla (Figures 3L–3U), tonic Mi4, Mi9, and Tm9 showed similar tuning as L1–L5 and again little surround dependency. However, for all transient cells (Mi1 and Tm3 for ON; Tm1, Tm2, and Tm4 for OFF), increasing background contrast had a strongly suppressive effect, which is a hallmark of divisive contrast normalization [41].

As with the corresponding behavioral experiments (Figure 2), linear receptive fields could not explain the effect given that the average luminance was constant for all conditions. Curves were shifted rightward on the logarithmic axis, which corresponds to divisive stretching in linear contrast space. Importantly, preferred direction responses in T4 and T5 were also strongly background dependent (Figures 3P and 3U) even

though not all their medulla inputs are subject to gain control. Finally, sensitivity to foreground contrast was generally higher in ON than OFF units.

Several cell types—particularly medulla transient cells—showed a dependency between fluorescence modulation at the target frequency and average response (Figure S4), possibly due to temporal integration by the calcium indicator [47]. Depending on this average activity, a saturating transformation between calcium signal and GCaMP fluorescence could by itself introduce compression of strong signal amplitudes due to ceiling effects at the far end of the sensor's dynamic range. To rule this out, we directly compared mean activity with oscillation amplitude and found no region in which this correlation was negative (Figures S4Q–S4S).

To quantify tuning curves in detail, we fit a closed-form model resembling common models of divisive normalization to the data (Figure 3V; STAR Methods) [41, 48]. Here, response gain is regulated by a divisive term that depends on background contrast while a linear term represents the combined contribution of foreground contrast and background leakage. The model accurately reproduced tuning curves for each cell type (Figure 3W; Table S1). Critically, it accounted for vertical shifts as well as sigmoidal tuning curves and context-dependent changes in contrast sensitivity.

We computed a normalization index from model parameters that estimates the degree of normalization. Given that different cell types had different baseline sensitivities and that horizontal shifts on a logarithmic scale correspond to multiplication, we quantified the relative factor by which tuning curves would shift when background contrast was increased from 0% to 100% (STAR Methods). This index was substantially higher in transient medulla cells (Mi1, Tm3, Tm1, Tm2, and Tm4) and direction-selective T4 and T5 cells than in L1–L5 or tonic medulla units (Mi4, Mi9, and Tm9; Figure 3X). Interestingly, L2 and L5 exhibited mildly elevated normalization indices. For L2, this may be related to previously described non-linearities in its receptive field structure [49].

Normalization Relies on Fast Integration of a Pool of Transient Units

Overall, fly contrast gain control appeared to be based on divisive normalization that predominantly originates in medulla units with transient response dynamics. We focused on these neurons to investigate the mechanism in detail. Responses in Mi1, Tm1, Tm2, and Tm3 were equally suppressed for all background grating directions relative to a reference stimulus with zero background contrast (Figure 4A). Temporal frequency tunings for suppression resembled band-pass filters with a peak at 2 Hz (Figure 4B). Crucially, static backgrounds did not have a suppressive effect. Suppression steadily increased with the outer diameter of an annulus containing the background pattern, which again indicated an extended integration area (Figure 4C). Spatiotemporal features of neural gain control thus matched our findings from behavior (Figure 2).

To determine the temporal scale of normalization, we designed a contrast-step stimulus in which the foreground was replaced by a single light pulse matching each cell type's polarity (Figure 4D). By varying the time interval between motion onset of the background grating and the onset of the pulse, we scanned

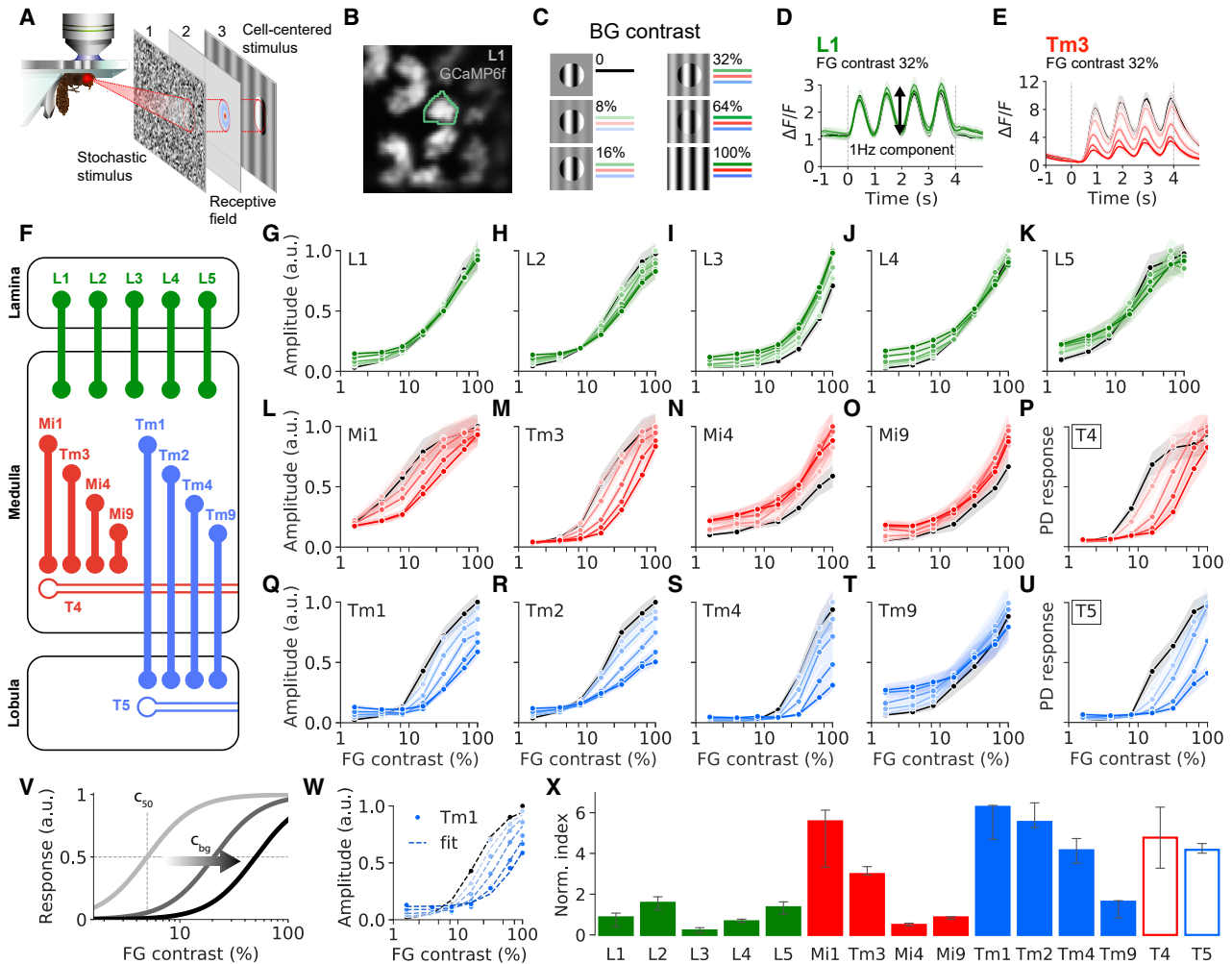


Figure 3. Contrast Normalization Emerges in Transient Medulla Neurons

(A) Schematic of experimental procedure. (1) White noise stimulus. (2) Receptive field reconstruction from single-neuron calcium signals. (3) Drifting grating with different contrasts in foreground and background.

(B) Two-photon image of L1 axon terminals expressing GCaMP6f. Green line indicates example region of interest.

(C) Experimental protocol. Darker color shade corresponds to higher background contrast as used in (G)–(U). Zero background contrast condition is shown in black.

(D and E) Average calcium responses of L1 (D) and Tm3 (E) for fixed foreground and various background contrasts.

(F) Schematic of the motion circuit including all neurons measured.

(G–K) Contrast tuning curves measured as amplitude of calcium signals at stimulus frequency for L1–L5. Shaded areas show bootstrapped 68% confidence intervals around the mean (L1 in G: 21/7 cells/flyes, L2 in H: 26/8, L3 in I: 23/6, L4 in J: 19/6, L5 in K: 18/9).

(L–P) Contrast tuning curves for ON pathway neurons (Mi1 in L: 20/5, Tm3 in M: 21/8, Mi4 in N: 20/13, Mi9 in O: 21/9, T4 in P: 23/10).

(Q–U) Contrast tuning curves for OFF pathway neurons (Tm1 in Q: 21/7, Tm2 in R: 20/6, Tm4 in S: 20/13, Tm9 in T: 19/6, T5 in U: 21/9).

(V) Illustration of divisive normalization model for tuning curves. Increasing background contrast c_{bg} shifts the sigmoidal tuning curve from baseline sensitivity c_{50} to higher contrasts.

(W) Example fit of model for Tm1.

(X) Normalization index for all neurons shown as median with 68% bootstrapped confidence intervals. Transient medulla neurons Mi1, Tm3, Tm1, Tm2, and Tm4, as well as T4 and T5, exhibited strongest degree of normalization.

See also [Figures S3 and S4](#), [Tables S1 and S2](#), and [Video S3](#).

the temporal profile of the suppressive signal. For the tested neurons Tm3 and Tm2, we found virtually immediate response reduction within a measurement precision of 50 ms given by the smallest tested onset difference. We observed transient ringing of suppression strength at the background temporal frequency. Ringing was stronger when the grating was present

before motion onset compared to when it was masked by uniform gray. A similar effect has been described in LPTCs [26], where it results from neural integration of multiple transient, out-of-phase inputs. In sum, these findings indicated that surround suppression derives from a pool of transient neurons that are not selective for direction. Both isotropy and frequency

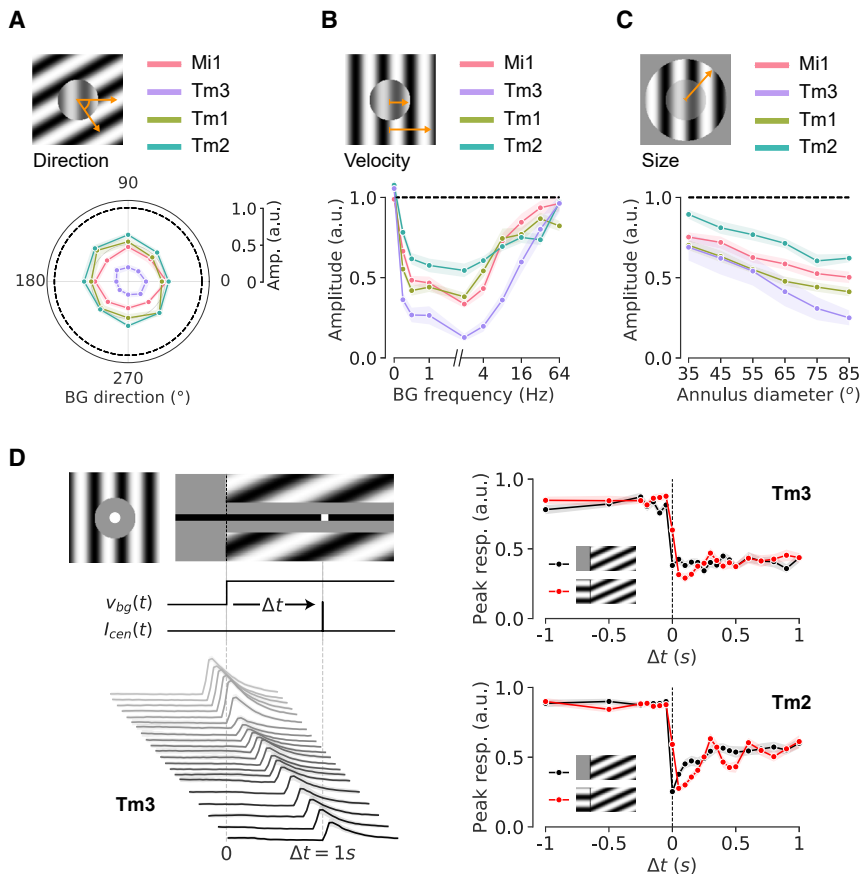


Figure 4. Neural Contrast Normalization Relies on Rapid Integration of a Pool of Transient Units

(A) Polar plot of response amplitude for different directions of background motion. Black dashed line represents response to reference stimulus with background contrast of 0%. For each neuron, foreground contrast was chosen to maximize possible background suppression (Mi1: 16%, Tm3: 32%, Tm1: 64%, Tm2: 100%).

(B) Responses for different background contrast frequencies, revealing band-pass tuning of suppression.

(C) Suppression strength increased with outer diameter of background annulus (Mi1: 21/9 cells/flyes, Tm3: 20/6, Tm1: 18/6, Tm2: 21/4 in A–C).

(D) Top left: x-y and x-t plots of contrast-step stimulus for Tm3 (ON center). Background contrast frequency was 3 Hz. Center left: velocity function $v_{bg}(t)$ of background and intensity function $I_{cen}(t)$ of center pulse. Bottom left: mean responses of Tm3 for different time intervals Δt . Right: mean peak amplitude for Tm3 and Tm2 (Tm3: 19/6, Tm2: 20/5). Black line shows condition where the background grating was masked before onset; red where background was visible but static.

Shaded areas around curves indicate bootstrapped 68% confidence intervals. See also modeling in Figure S5, Table S2, and Video S4.

tunings were strikingly similar to filter properties of the transient lamina and medulla units involved in motion detection (Figures S3U and S3V). This suggested that one or more of these cell types provides input to the suppressive pool.

To determine whether a mechanism that integrates transient units across space to divisively suppress local responses could reproduce our findings, we built a time-resolved, data-driven model. The model faithfully predicted direction, frequency, and size tunings, as well as contrast-step ringing, T4 and T5 responses, and LPTC output for our behavioral stimuli (Figure S5A–K).

Neural Feedback Is Critical for Contrast Normalization

Spatial pooling, however, could occur over either feedforward signals from the lamina or feedback from the medulla (Figure 5A). In vertebrate systems, it has proven difficult to distinguish the two [41, 50, 51]. Fly transient units in the lamina or medulla have similar temporal properties (Figures S3U and S3V), and both implementations produce equivalent steady-state output [48], so we used genetic silencing to pinpoint the source. We co-expressed a calcium indicator and the tetanus toxin light chain (TNT; STAR Methods) [52] in different medulla cell types, blocking chemical synaptic output and thus feedback from the entire neuron array but leaving feedforward input and calcium signals intact.

For the ON pathway unit Tm3, we observed significantly reduced suppression across background frequencies when compared to controls with inactive TNT (Figures 5B and 5C).

constant of model fits was significantly increased (Figure 5D). This suggests that Tm3 cells were disinhibited due to a reduced pool signal. We observed similar effects for ON-sensitive Mi1 cells (Figures 5E and 5F), but the impact was less pronounced than for Tm3 cells. Absolute signal amplitude was generally not affected by silencing, demonstrating that cells remained visually responsive in the presence of TNT (see Figure 5B).

In the OFF pathway, blocking Tm1 cells did not have any significant effects (Figures 5G and 5H). In contrast, when blocking Tm2, we observed an almost complete loss of background suppression across frequencies (Figure 5I). For this cell type, we did not observe any change in contrast tuning curves for the 0% background condition, and consequently, the fitted semi-saturation constant was not affected (Figure 5J). For full background contrast, however, suppression at high foreground contrasts was strongly reduced. Additionally, background leakage at low foreground contrasts increased substantially compared to control flies. As with Tm3 and Mi1, this is compatible with Tm2 cells being disinhibited due to the silencing of a suppressive signal derived from recurrent output. We therefore conclude that in the fly, contrast normalization is at least partially based on feedback from a combination of medulla neurons.

Contrast Normalization Improves Robustness to Natural Scene Variability

Could this type of response normalization account for the robustness of fly motion detection? Previous work on EMDs

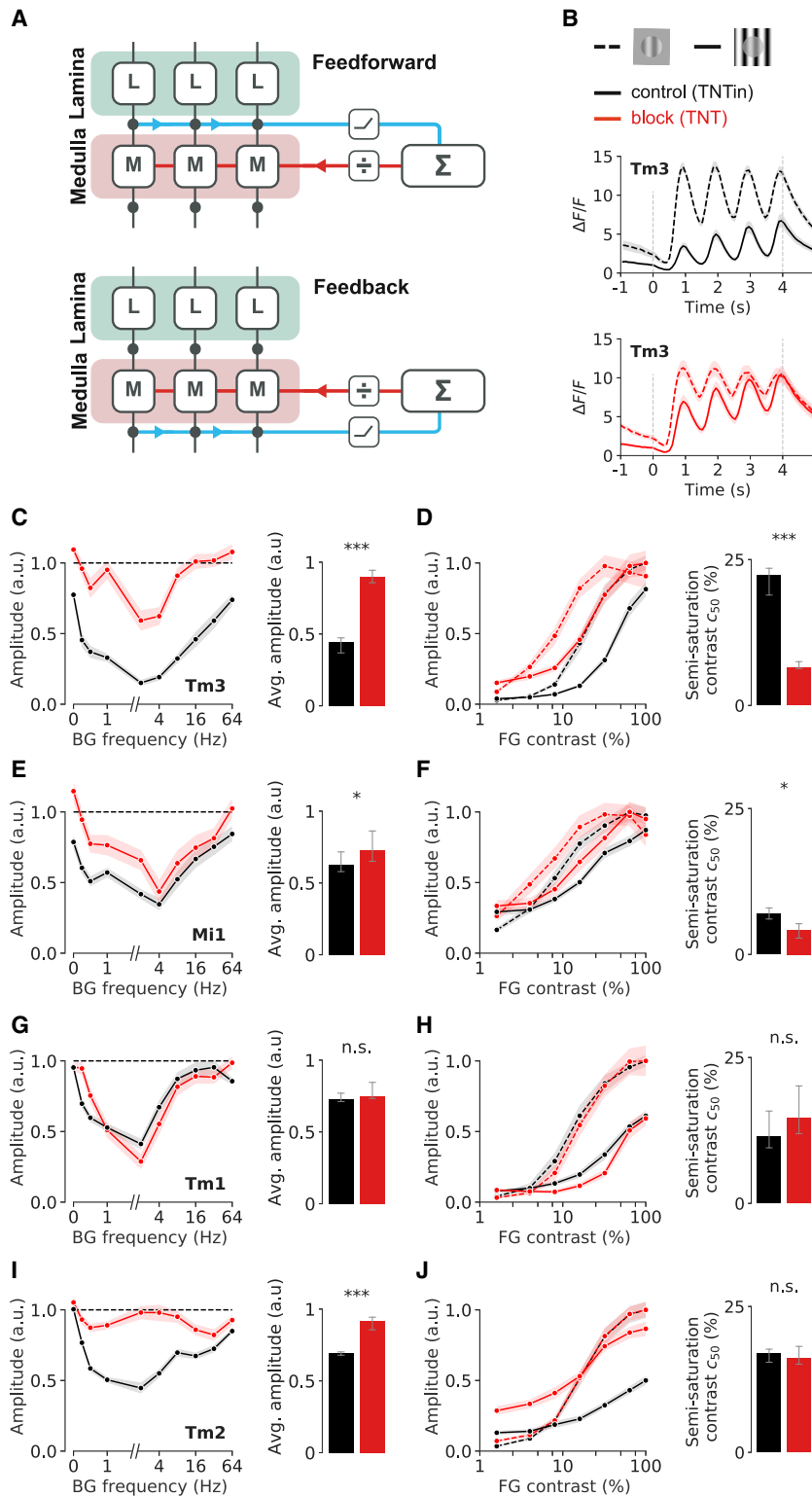


Figure 5. Neural Feedback Underlies Contrast Normalization

(A) Schematic of feedforward and feedback model for surround suppression.

(B) Mean responses of Tm3 for TNT block (red) and TNTIn controls (black) at background frequency 16 Hz (dashed line indicates reference response and solid line the response at full background contrast; Tm3 block: 21/5 cells/flyes, Tm3 control: 20/5).

(C) Left: frequency tuning for block experiment. Black dashed line represents response to reference stimulus. Right: average amplitude over all frequencies was higher for Tm3 block flies (Mann-Whitney U: 8, *** $p < 0.001$).

(D) Left: foreground contrast tuning for block experiments at 0% and 100% background contrast. Right: contrast sensitivity was increased for Tm3 block flies as measured by lowered semi-saturation constant c_{50} (Mann-Whitney U: 39, *** $p < 0.001$).

(E) Blocking results for Mi1 (as in C). Average amplitude over all frequencies was reduced for Mi1 block flies (Mi1 block: 20/5, Mi1 control: 21/6; Mann-Whitney U: 143, * $p = 0.04$).

(F) Blocking results for Mi1 (as in D). Contrast sensitivity was increased for Mi1 block flies (Mann-Whitney U: 128, * $p = 0.02$).

(G) Blocking results for Tm1 (as in C). No significant effect was found for Tm1 block flies (Tm1 block: 20/5, Tm1 control: 19/5; Mann-Whitney U: 169, NS $p = 0.28$).

(H) Blocking results for Tm1 (as in D). Sensitivity was not affected (Mann-Whitney U: 158, NS $p = 0.19$).

(I) Blocking results for Tm2 (as in C; Tm2 block: 20/5, Tm2 control: 25/6; Mann-Whitney U: 17, *** $p < 0.001$).

(J) Blocking results for Tm2 (as in D; Mann-Whitney U: 239, NS $p = 0.49$). Semi-saturation constant at 0% background contrast did not change for Tm2 block flies. Shaded areas show bootstrapped 68% confidence intervals around the mean. Error bars show bootstrapped 68% confidence intervals around the median.

See also Table S2.

data-driven LPTC model and found moderate reduction of cross-image variability compared to a model with bypassed normalization (Figures S5L–S5N). However, post hoc ablation may specifically disadvantage the simpler model. To investigate performance limits in a principled way, we pursued a task-driven approach.

Recent progress in deep artificial networks has made it feasible to use image-processing models of neural systems for rigorously assessing performance on real-world problems [54–56]. EMD-like architectures are concisely expressed as multi-layer convolutional networks [54] and fully differentiable, rendering them amenable to optimization methods like gradient descent.

and natural scenes has exploited compressive transforms but did so heuristically or without surround-dependent gain control [42, 43, 53]. We evaluated natural image responses in the

We designed a fly-like neural network and independently trained possible types of contrast processing such that each model class could optimally adapt to a specific, behaviorally relevant task.

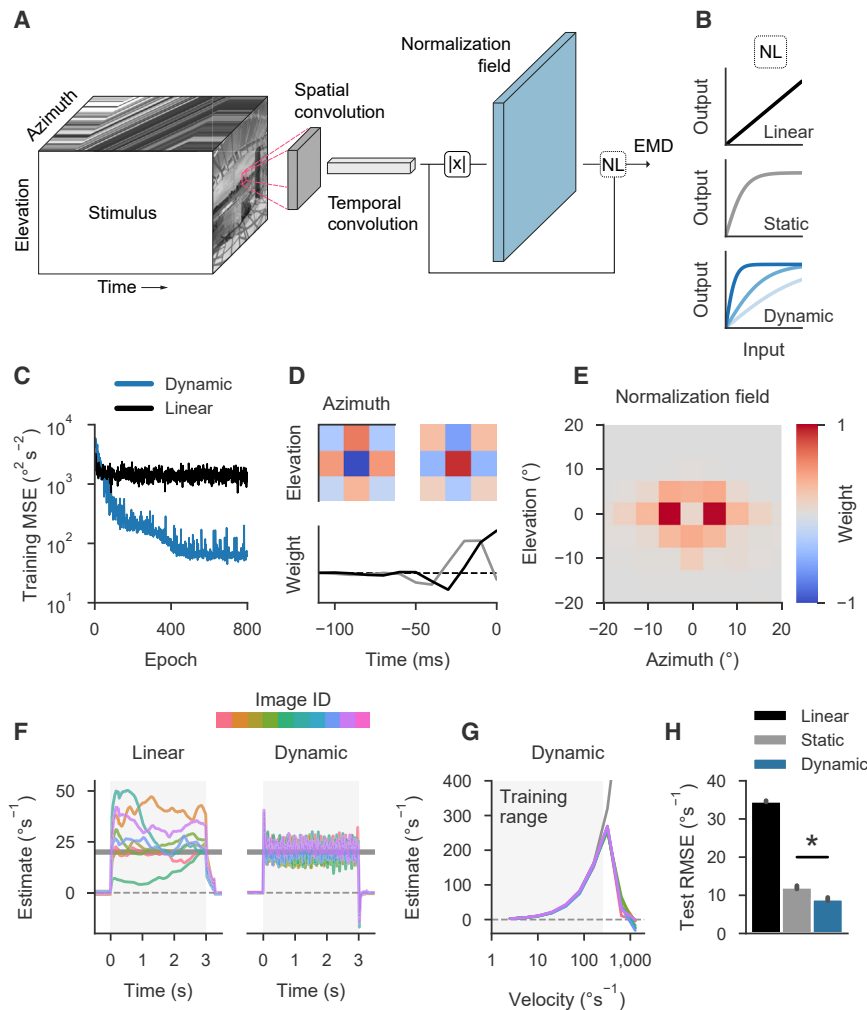


Figure 6. Contrast Normalization Enhances Robustness to Natural Scene Variability

(A) Schematic of single convolutional input filter. Motion stimuli are sequentially processed by a spatial $3 \times 3 \times 1$ (azimuth, elevation, time) and a temporal $1 \times 1 \times 30$ filter. Through a transfer function, the signal is combined with a normalization signal generated by a $11 \times 11 \times 1$ convolution operating on full-wave rectified input signal. The output of two distinct channels is processed analogously to multiplicative EMDs.

(B) Input-output relationships for linear, static, and dynamic models. In the dynamic model, response sensitivity is a function of normalization field activity. (C) Training mean squared error (MSE) for two example models during stochastic gradient descent.

(D) Spatial and temporal receptive fields for the two channels of a typical dynamic model. Depicted are normalized filter weights. (E) Spatial receptive field of normalization pool for the model from (D).

(F) Model output for individual images moving at $20^\circ/\text{s}$ during gray-shaded period. Gray line indicates target velocity. Left: example model without non-linearity. Right: example model with dynamic non-linearity.

(G) Velocity tuning curves of example dynamic model for individual images (averaged between 0 and 3 s after motion onset). Gray line indicates true velocity. Gray-shaded area indicates the 99th percentile of absolute velocities in training set.

(H) Mean performance of trained models on held-out test set, estimated as root mean square error (RMSE; $n = 22/23/16$ for linear/static/dynamic; $*p < 0.001$, $t = 9.01$, Student's t test with assumed equal variance; only difference between static and dynamic was tested). Error bars indicate bootstrapped 68% confidence intervals. See also Figure S6.

All models featured linear, spatiotemporally separable input convolutions (Figure 6A). We evaluated three alternatives for contrast transformation: a linear stage where output was transmitted unchanged, a statically compressive stage that limited signal range independently of context, and a dynamic compression stage with adaptive gain depending on the output of a contrast-sensitive surround filter (Figures 6A and 6B; STAR Methods). Resulting output from two distinct channels was then processed according to a multiplicative EMD scheme. Through backpropagation and stochastic gradient descent, models were trained to estimate the true velocity of natural images translating at random speeds.

All models successfully learned the task on the training set (Figure 6C). We initialized convolutions randomly but after training observed antagonistic spatial filters and transient temporal filters where one channel was phase delayed with respect to the other (Figure 6D; Figures S6A–S6C). Models thus made extensive use of redundancy reduction through center-surround configurations [35] and discovered the EMD strategy of delay and compare [26]. Normalization fields for the dynamic model spanned approximately 30° in azimuth and invariably excluded information from the center of the filter (Figure 6E; Figure S6C). Interestingly, dynamic models exploited normalization in both

channels and switched normalization strategies during training, transitioning from purely static to purely context-dependent compression (Figures S6D and S6E). Overall, normalized networks acquired representations that matched filtering and gain control properties of the fly medulla.

When tested on previous experimental stimuli (Figure 1), linear models exhibited improved velocity tuning curves compared to a standard EMD (Figures 1F and 6F; Figure S6F), but estimates still varied substantially across time. Dynamic models, on the other hand, proved extremely robust at extracting scene motion across time, images, and velocities within the velocity range of the training set (Figures 6F and 6G). Given that all networks were based on a multiplicative EMD scheme, typical phenomena like the velocity optimum were still present. We compared average estimation error on a held-out test set and found both types of non-linear compression to vastly outperform the linear stage (Figure 6H). The performance of static compression indicates that simple response saturation already enhances robustness to contrast fluctuations in natural scenes. However, fly-like context sensitivity consistently decreased test error over the static non-linearity (error reduction 22.0%–29.2%; bootstrapped 95% CI). Finally, we benchmarked generalization on a fully independent image set (Figure S6G), where linear models failed

catastrophically while both compressive stages retained performance. This was particularly pronounced when testing images with high dynamic range (STAR Methods). Critically, on all datasets, dynamic compression resulted in substantial error reduction with respect to both linear transfer and static compression.

DISCUSSION

In summary, our work represents the first demonstration that divisive contrast normalization occurs in the fly visual system and offers a comprehensive look at non-linear response properties in a virtually complete motion vision circuit. We established at multiple levels of motion processing that responses to moving panoramas are substantially more robust than predicted by correlation-based models of the system. Our behavioral experiments indicate that the sensitivity of the optomotor response is regulated by average contrast in a spatially confined part of the visual field. Critically, we traced the emergence of this dynamic signal compression to local elements in the medulla of the fly optic lobe and used targeted circuit manipulation to identify neural feedback as a critical underlying mechanism. Finally, our task-driven approach revealed that the inclusion of spatial contrast normalization drastically improves velocity estimation in correlation-based models of fly motion vision.

Implications for Fly Motion Vision

Previous work on the function of local units in the *Drosophila* optic lobe mostly explored linear properties of light responses, often relying on first-order systems identification techniques like reverse correlation [9–11, 29]. Investigation of non-linear contributions generally focused on computations in direction-selective T4 and T5 cells [13–19, 57, 58].

Here, we describe a powerful non-linearity, adaptive gain control that occurs in a majority of columnar neurons involved in the detection of motion. This casts doubt on the extent to which existing functional descriptions can be generalized. Linear filter estimates are typically based on responses to dynamic noise stimuli of fixed amplitude [9, 10, 29]. Our work suggests that this contrast regime only corresponds to one particular adaptation state for any measured cell type, so filter properties may well differ for stimuli with differing contrast characteristics. Step and edge responses, for instance, are usually measured on backgrounds with uniform luminance [11, 12, 18, 59]. This places cells in a maximally sensitive state due to lack of surround inhibition and is likely to affect both response amplitude and kinetics. Signal compression may reconcile observed discrepancies between studies conducted with different stimuli.

Interestingly, visual interneurons exhibited qualitatively different sensitivity curves even at constant background contrast. In the lamina, for instance, only tonic cell type L3 responded linearly to increasing visual contrast. Sensitivity curves of transiently responding cell types like L1 and L2, on the other hand, proved approximately logarithmic. This is in line with expectations from previous work in other fly species [60] but deviates from predictions based on white noise characterizations [29].

Moreover, we observed a stark discrepancy in baseline sensitivity between ON- and OFF-sensitive neurons, where tuning curves of dark-selective units were shifted toward higher pattern contrast. Notably, due to strong surround suppression, full-field

gratings elicited comparatively weak responses in T5 units whereas T4 cells were driven effectively by the same stimuli. This adds to previous work on ON-OFF asymmetries in the *Drosophila* visual system [31, 38]. We conclude that even at primary processing stages, the fly visual system represents contrast in a multiplexed fashion where individual channels diverge with respect to how they transmit information about luminance differences. The function of these asymmetries remains to be investigated.

The proposed model based on divisive normalization accurately captures most features of the observed contrast tuning curves (see Figure 3; Figure S5; Table S1). Certain discrepancies remain. For instance, the normalization model predicts that responses for different background contrasts eventually plateau at the same level. However, we observed in both behavior (Figure 2E) and T4 responses (Figure 3P) that in the absence of background contrast, saturation occurred at a lower level than for other conditions. To explain such non-monotonic behavior, further investigation of the underlying mechanism is required.

Divisive normalization of local motion signals has previously been suggested to occur at the level of LPTCs, through either isotropic pooling of EMDs in hypothetical secondary cell types [33] or passive membrane properties of LPTCs [61, 62]. Here, we show that gain control already originates upstream of motion-sensitive cell types T4 and T5. However, LPTC-intrinsic gain control mechanisms, including temporal adaptation [32], could well be complementary such that at each processing stage, the fly visual system makes use of compression to optimize the reliability of output signals.

In flies, there is ample evidence for changes in visual coding that depend on the behavioral state of the animal. Various interneurons within the optic lobe, for instance, are affected by the activity of octopaminergic projection units, leading to drastic shifts in response gain or temporal tuning [9, 59, 63–67]. Our calcium imaging experiments were performed in immobilized *Drosophila*. It will be of interest to explore whether the properties of contrast gain control are modulated by locomotion, particularly in highly state-sensitive units like Mi4 [59].

Mechanism of Signal Compression

Our experiments suggest that neural feedback plays a crucial role in gain adjustment. At this point, the cellular origin of feedback is unknown. Present experiments indicate a visual integration field that spans many columns (Figures 2 and 4). Moreover, the observed contrast compression appears to be suppressive. All tested medulla cell types with strong background contrast dependency emit acetylcholine, which, in the *Drosophila* visual system, is generally thought to be excitatory [68, 69]. Inhibitory interneurons could mediate the required synaptic sign reversal. Signal compression could then be implemented through lateral neighbor-to-neighbor interactions between columnar medulla units where suppressive signals spread through a local network. Alternatively, we hypothesize that wide-field interneurons pool local medulla units across multiple columns and provide recurrent inhibitory input to the same cells. In our data-driven model, such a pool cell mechanism accounted for all observed spatio-temporal properties of signal compression including ringing effects (Figure S5). Finally, our TNT-based intervention strategy

should leave coupling via electrical synapses intact [52, 70]. We can therefore not exclude that gap junctions are also involved in shaping contrast response properties.

Silencing feedback from individual medulla cell types had differential effects, ranging from completely abolished suppression in Tm2 to unchanged responses in Tm1 (Figure 5). This suggests either that multiple cell types feed into the pool signal with varying weight or that alternative mechanisms provide the compressive signal, for example, in Tm1. Moreover, it is an open question whether all cell types are suppressed by one or multiple pool cell types. Asymmetries in sensitivity between ON and OFF pathways, for instance, could be an indicator for polarity-specific sources of suppression.

In both distal and proximal layers of the medulla, the class of neuropil-intrinsic Dm and Pm neurons contains approximately 20 cell types and offers a possible substrate for the mechanism [71, 72]. These neurons arborize within the medulla and exhibit diverse stratification and tiling patterns, often spanning dozens of columns and thus approximately matching the observed suppression field of local units. Dm and Pm units release either GABA or glutamate for which receptors in the fly visual system are mostly inhibitory [68], pointing to these cell types as potential candidates for gain control.

Functional Relevance

Normalization has often been described as a generic mechanism for removing higher-order correlations from natural signals [5, 73–75]. Here, we close the loop between neural mechanism and an ecologically critical behavior, the optomotor response, and demonstrate how contrast gain control can render motion detection resilient to challenges imposed by natural scene statistics. Specifically, normalization serves to distinguish between ecologically relevant parameters like retinal image velocity and nuisance factors like image contrast.

Various biomimetic modeling studies have incorporated compressive transforms along the motion processing cascade to improve robustness under naturalistic visual conditions [43, 53, 76]. In contrast to our work, these normalization stages were not based on experimental evidence, required *ad hoc* parameter tuning, and generally operated in the temporal domain. Interestingly, the fly visual system bases gain control on a temporally immediate, spatially extended estimate of contrast. This represents a trade-off where spatial resolution is sacrificed in favor of temporal resolution, which may be advantageous for global optic flow estimation in rapidly moving animals.

To assess the exact causal contribution of contrast compression to the robustness of velocity estimation in *Drosophila*, one would need to disrupt this mechanism specifically while leaving all other visual processing intact. Silencing the synaptic output of medulla neurons (Figure 5) demonstrates the importance of neural feedback for gain control but should additionally affect feedforward processing in downstream units, particularly T4 and T5 [11, 12, 77, 78]. Future mapping of the circuits underlying contrast compression will provide the tools for establishing causality.

The convolutional network (Figure 6) solves the task of estimating velocity across diverse environments and at little computational cost, particularly compared to standard optic flow algorithms like the Lucas-Kanade method [79]. Present findings

may thus aid the design of low-power, low-latency machine vision systems suitable for autonomous vehicles [80, 81].

Comparison with Other Sensory Systems

Gain control in the *Drosophila* optic lobe bears a striking resemblance to normalization in other systems and modalities like fly olfaction [82] or mammalian auditory cortex [83] as well as processing in vertebrate visual areas from retina to V1 [48, 84–86]. Spatial and temporal tuning or isotropy of non-linear surround suppression in the lateral geniculate nucleus, in particular, qualitatively match that of transient units in the fly medulla [40]. The present study suggests differences at the implementation level. For instance, investigations into divisive normalization in mammalian V1 cells point to feedforward mechanisms underlying gain control whereas the fly visual system appears to rely primarily on feedback signals (Figure 5) [50]. Both systems, however, realize a similar algorithm. This provides further proof for evolutionary convergence on canonical solutions in neural sensory processing [41].

Overall, our work establishes the *Drosophila* visual system with its defined cell types, known connectivity patterns, powerful genetic toolkit, and direct correspondence between circuit and task as a novel model for the study of normalization. It thus lays the foundation for future mechanistic inquiries into the functional, cellular, molecular, and biophysical underpinnings of a crucial computation in sensory processing.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Natural image sets
 - Behavioral experiments
 - Electrophysiology
 - Calcium imaging
 - Modeling
- QUANTIFICATION AND STATISTICAL ANALYSIS
 - Data evaluation for behavioral experiments
 - Data evaluation for electrophysiological experiments
 - Data evaluation for calcium imaging experiments
 - Statistical tests
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cub.2019.10.035>.

ACKNOWLEDGMENTS

We thank S. Prech for technical assistance; J. Kuhl for schematics; A. Nern, G.M. Rubin, and T. Schilling for supplying unpublished Gal4 driver lines; G. Ammer for immunohistochemistry; R. Brinkworth for providing access to natural images; and L. Groschner, A. Mauss, and M. Meier for commenting on drafts of the manuscript. The study was supported by the DFG (SFB 870) and the Max Planck Society.

AUTHOR CONTRIBUTIONS

M.S.D., A.L., and A.B. jointly conceived the study. M.S.D. and A.L. designed all experiments. A.L. and L.B. conducted behavioral experiments. E.S. recorded electrophysiological responses. M.S.D., N.P., F.G.R., and A.S. performed calcium imaging. A.L. designed and analyzed the convolutional model. M.S.D. and A.L. analyzed data, performed modeling, and wrote the manuscript. All authors participated in editing the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 2, 2019

Revised: September 17, 2019

Accepted: October 18, 2019

Published: January 9, 2020

REFERENCES

- Gibson, J.J. (1950). *The Perception of the Visual World* (Houghton Mifflin).
- Dickinson, M.H. (2014). Death Valley, *Drosophila*, and the Devonian toolkit. *Annu. Rev. Entomol.* *59*, 51–72.
- Borst, A. (2014). Fly visual course control: behaviour, algorithms and circuits. *Nat. Rev. Neurosci.* *15*, 590–599.
- Geisler, W.S. (2008). Visual perception and the statistical properties of natural scenes. *Annu. Rev. Psychol.* *59*, 167–192.
- Rieke, F., and Rudd, M.E. (2009). The challenges natural images pose for visual adaptation. *Neuron* *64*, 605–616.
- Mauss, A.S., Vlasits, A., Borst, A., and Feller, M. (2017). Visual circuits for direction selectivity. *Annu. Rev. Neurosci.* *40*, 211–230.
- Yang, H.H., and Clandinin, T.R. (2018). Elementary motion detection in *Drosophila*: Algorithms and mechanisms. *Annu. Rev. Vis. Sci.* *4*, 143–163.
- Shinomiya, K., Huang, G., Lu, Z., Parag, T., Xu, C.S., Aniceto, R., Ansari, N., Cheatham, N., Lauchie, S., Neace, E., et al. (2019). Comparisons between the ON- and OFF-edge motion pathways in the *Drosophila* brain. *eLife* *8*, e40025.
- Arenz, A., Drews, M.S., Richter, F.G., Ammer, G., and Borst, A. (2017). The temporal tuning of the *Drosophila* motion detectors is determined by the dynamics of their input elements. *Curr. Biol.* *27*, 929–944.
- Behnia, R., Clark, D.A., Carter, A.G., Clandinin, T.R., and Desplan, C. (2014). Processing properties of ON and OFF pathways for *Drosophila* motion detection. *Nature* *512*, 427–430.
- Serbe, E., Meier, M., Leonhardt, A., and Borst, A. (2016). Comprehensive characterization of the major presynaptic elements to the *Drosophila* OFF motion detector. *Neuron* *89*, 829–841.
- Strother, J.A., Wu, S.-T., Wong, A.M., Nern, A., Rogers, E.M., Le, J.Q., Rubin, G.M., and Reiser, M.B. (2017). The emergence of directional selectivity in the visual motion pathway of *Drosophila*. *Neuron* *94*, 168–182.e10.
- Salazar-Gatzimas, E., Agrochao, M., Fitzgerald, J.E., and Clark, D.A. (2018). The neuronal basis of an illusory motion percept is explained by decorrelation of parallel motion pathways. *Curr. Biol.* *28*, 3748–3762.e8.
- Maisak, M.S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., Schilling, T., Bahl, A., Rubin, G.M., Nern, A., et al. (2013). A directional tuning map of *Drosophila* elementary motion detectors. *Nature* *500*, 212–216.
- Gruntman, E., Romani, S., and Reiser, M.B. (2018). Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in *Drosophila*. *Nat. Neurosci.* *21*, 250–257.
- Salazar-Gatzimas, E., Chen, J., Creamer, M.S., Mano, O., Mandel, H.B., Matulis, C.A., Pottackal, J., and Clark, D.A. (2016). Direct measurement of correlation responses in *Drosophila* elementary motion detectors reveals fast timescale tuning. *Neuron* *92*, 227–239.
- Leong, J.C.S., Esch, J.J., Poole, B., Ganguli, S., and Clandinin, T.R. (2016). Direction selectivity in *Drosophila* emerges from preferred-direction enhancement and null-direction suppression. *J. Neurosci.* *36*, 8078–8092.
- Haag, J., Arenz, A., Serbe, E., Gabbiani, F., and Borst, A. (2016). Complementary mechanisms create direction selectivity in the fly. *eLife* *5*, e17421.
- Fisher, Y.E., Silies, M., and Clandinin, T.R. (2015). Orientation selectivity sharpens motion detection in *Drosophila*. *Neuron* *88*, 390–402.
- Bahl, A., Ammer, G., Schilling, T., and Borst, A. (2013). Object tracking in motion-blind flies. *Nat. Neurosci.* *16*, 730–738.
- Busch, C., Borst, A., and Mauss, A.S. (2018). Bi-directional control of walking behavior by horizontal optic flow sensors. *Curr. Biol.* *28*, 4037–4045.e5.
- Haikala, V., Joesch, M., Borst, A., and Mauss, A.S. (2013). Optogenetic control of fly optomotor responses. *J. Neurosci.* *33*, 13927–13934.
- Krapp, H.G., and Hengstenberg, R. (1996). Estimation of self-motion by optic flow processing in single visual interneurons. *Nature* *384*, 463–466.
- Hassenstein, B., and Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*. *Z. Naturforsch. B 11*, 513–524.
- Joesch, M., Plett, J., Borst, A., and Reiff, D.F. (2008). Response properties of motion-sensitive visual interneurons in the lobula plate of *Drosophila melanogaster*. *Curr. Biol.* *18*, 368–374.
- Borst, A., Reisenman, C., and Haag, J. (2003). Adaptation of response transients in fly motion vision. II: Model studies. *Vision Res.* *43*, 1309–1322.
- Borst, A., Flanagan, V.L., and Sompolinsky, H. (2005). Adaptation without parameter change: Dynamic gain control in motion detection. *Proc. Natl. Acad. Sci. USA* *102*, 6172–6176.
- Eichner, H., Joesch, M., Schnell, B., Reiff, D.F., and Borst, A. (2011). Internal structure of the fly elementary motion detector. *Neuron* *70*, 1155–1164.
- Clark, D.A., Bursztyn, L., Horowitz, M.A., Schnitzer, M.J., and Clandinin, T.R. (2011). Defining the computational structure of the motion detector in *Drosophila*. *Neuron* *70*, 1165–1177.
- Straw, A.D., Rainsford, T., and O'Carroll, D.C. (2008). Contrast sensitivity of insect motion detectors to natural images. *J. Vis.* *8*, 1–9.
- Leonhardt, A., Ammer, G., Meier, M., Serbe, E., Bahl, A., and Borst, A. (2016). Asymmetry of *Drosophila* ON and OFF motion detectors enhances real-world velocity estimation. *Nat. Neurosci.* *19*, 706–715.
- Harris, R.A., O'Carroll, D.C., and Laughlin, S.B. (2000). Contrast gain reduction in fly motion adaptation. *Neuron* *28*, 595–606.
- Reichardt, W., Poggio, T., and Hausen, K. (1983). Figure-ground discrimination by relative movement in the visual system of the fly. *Biol. Cybern.* *46*, 1–30.
- van Hateren, J.H. (1997). Processing of natural time series of intensities by the visual system of the blowfly. *Vision Res.* *37*, 3407–3416.
- Srinivasan, M.V., Laughlin, S.B., and Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proc. R. Soc. Lond. B Biol. Sci.* *216*, 427–459.
- Mauss, A.S., Pankova, K., Arenz, A., Nern, A., Rubin, G.M., and Borst, A. (2015). Neural circuit to integrate opposing motions in the visual field. *Cell* *162*, 351–362.
- Fitzgerald, J.E., and Clark, D.A. (2015). Nonlinear circuits for naturalistic visual motion estimation. *eLife* *4*, e09123.
- Clark, D.A., Fitzgerald, J.E., Ales, J.M., Gohl, D.M., Silies, M.A., Norcia, A.M., and Clandinin, T.R. (2014). Flies and humans share a motion estimation strategy that exploits natural scene statistics. *Nat. Neurosci.* *17*, 296–303.
- Demb, J.B. (2008). Functional circuitry of visual adaptation in the retina. *J. Physiol.* *586*, 4377–4384.

40. Bonin, V., Mante, V., and Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus. *J. Neurosci.* *25*, 10844–10856.
41. Carandini, M., and Heeger, D.J. (2011). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* *13*, 51–62.
42. Dror, R.O., O'Carroll, D.C., and Laughlin, S.B. (2001). Accuracy of velocity estimation by Reichardt correlators. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* *18*, 241–252.
43. Brinkworth, R.S.A., and O'Carroll, D.C. (2009). Robust models for optic flow coding in natural scenes inspired by insect biology. *PLoS Comput. Biol.* *5*, e1000555.
44. Bahl, A., Serbe, E., Meier, M., Ammer, G., and Borst, A. (2015). Neural mechanisms for *Drosophila* contrast vision. *Neuron* *88*, 1240–1252.
45. Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., et al. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* *499*, 295–300.
46. Silies, M., Gohl, D.M., Fisher, Y.E., Freifeld, L., Clark, D.A., and Clandinin, T.R. (2013). Modular use of peripheral input channels tunes motion-detecting circuitry. *Neuron* *79*, 111–127.
47. Strother, J.A., Nern, A., and Reiser, M.B. (2014). Direct observation of ON and OFF pathways in the *Drosophila* visual system. *Curr. Biol.* *24*, 976–983.
48. Heeger, D.J. (1992). Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* *9*, 181–197.
49. Freifeld, L., Clark, D.A., Schnitzer, M.J., Horowitz, M.A., and Clandinin, T.R. (2013). GABAergic lateral interactions tune the early stages of visual processing in *Drosophila*. *Neuron* *78*, 1075–1089.
50. Freeman, T.C.B., Durand, S., Kiper, D.C., and Carandini, M. (2002). Suppression without inhibition in visual cortex. *Neuron* *35*, 759–771.
51. Li, B., Thompson, J.K., Duong, T., Peterson, M.R., and Freeman, R.D. (2006). Origins of cross-orientation suppression in the visual cortex. *J. Neurophysiol.* *96*, 1755–1764.
52. Sweeney, S.T., Brodie, K., Keane, J., Niemann, H., and O'Kane, C.J. (1995). Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* *14*, 341–351.
53. Shoemaker, P.A., O'Carroll, D.C., and Straw, A.D. (2005). Velocity constancy and models for wide-field visual motion detection in insects. *Biol. Cybern.* *93*, 275–287.
54. LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* *521*, 436–444.
55. Yamins, D.L.K., and DiCarlo, J.J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.* *19*, 356–365.
56. McIntosh, L., Maheswaranathan, N., Nayebi, A., Ganguli, S., and Baccus, S. (2016). Deep learning models of the retinal response to natural scenes. In *Advances in Neural Information Processing Systems* 29, D.D. Lee, M. Sugiyama, U.V. Luxburg, I. Guyon, and R. Garnett, eds., pp. 1369–1377.
57. Haag, J., Mishra, A., and Borst, A. (2017). A common directional tuning mechanism of *Drosophila* motion-sensing neurons in the ON and in the OFF pathway. *eLife* *6*, e29044.
58. Wienecke, C.F.R., Leong, J.C.S., and Clandinin, T.R. (2018). Linear summation underlies direction selectivity in *Drosophila*. *Neuron* *99*, 680–688.e4.
59. Strother, J.A., Wu, S.-T., Rogers, E.M., Eliason, J.L.M., Wong, A.M., Nern, A., and Reiser, M.B. (2018). Behavioral state modulates the ON visual motion pathway of *Drosophila*. *Proc. Natl. Acad. Sci. USA* *115*, E102–E111.
60. Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch., C, Biosci.* *36*, 910–912.
61. Borst, A., Egelhaaf, M., and Haag, J. (1995). Mechanisms of dendritic integration underling gain control in fly motion-sensitive interneurons. *J. Comput. Neurosci.* *2*, 5–18.
62. Weber, F., Machens, C.K., and Borst, A. (2010). Spatiotemporal response properties of optic-flow processing neurons. *Neuron* *67*, 629–642.
63. Suver, M.P., Mamiya, A., and Dickinson, M.H. (2012). Octopamine neurons mediate flight-induced modulation of visual processing in *Drosophila*. *Curr. Biol.* *22*, 2294–2302.
64. Tuthill, J.C., Nern, A., Rubin, G.M., and Reiser, M.B. (2014). Wide-field feedback neurons dynamically tune early visual processing. *Neuron* *82*, 887–895.
65. Jung, S.N., Borst, A., and Haag, J. (2011). Flight activity alters velocity tuning of fly motion-sensitive neurons. *J. Neurosci.* *31*, 9231–9237.
66. Maimon, G., Straw, A.D., and Dickinson, M.H. (2010). Active flight increases the gain of visual motion processing in *Drosophila*. *Nat. Neurosci.* *13*, 393–399.
67. Chiappe, M.E., Seelig, J.D., Reiser, M.B., and Jayaraman, V. (2010). Walking modulates speed sensitivity in *Drosophila* motion vision. *Curr. Biol.* *20*, 1470–1475.
68. Davis, F.P., Nern, A., Picard, S., Reiser, M.B., Rubin, G.M., Eddy, S.R., and Henry, G.L. (2018). A genetic, genomic, and computational resource for exploring neural circuit function. *bioRxiv*. <https://doi.org/10.1101/385476v1>.
69. Mauss, A.S., Meier, M., Serbe, E., and Borst, A. (2014). Optogenetic and pharmacologic dissection of feedforward inhibition in *Drosophila* motion vision. *J. Neurosci.* *34*, 2254–2263.
70. Joesch, M., Schnell, B., Raghu, S.V., Reiff, D.F., and Borst, A. (2010). ON and OFF pathways in *Drosophila* motion vision. *Nature* *468*, 300–304.
71. Nern, A., Pfeiffer, B.D., and Rubin, G.M. (2015). Optimized tools for multi-color stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proc. Natl. Acad. Sci. USA* *112*, E2967–E2976.
72. Fischbach, K.F., and Dittrich, A.P. (1989). The optic lobe of *Drosophila melanogaster*. I: A Golgi analysis of wild-type structure. *Cell Tissue Res.* *258*, 441–475.
73. Schwartz, O., and Simoncelli, E.P. (2001). Natural signal statistics and sensory gain control. *Nat. Neurosci.* *4*, 819–825.
74. Mante, V., Frazor, R.A., Bonin, V., Geisler, W.S., and Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nat. Neurosci.* *8*, 1690–1697.
75. Mante, V., Bonin, V., and Carandini, M. (2008). Functional mechanisms shaping lateral geniculate responses to artificial and natural stimuli. *Neuron* *58*, 625–638.
76. Shoemaker, P.A., Hyslop, A.M., and Humbert, J.S. (2011). Optic flow estimation on trajectories generated by bio-inspired closed-loop flight. *Biol. Cybern.* *104*, 339–350.
77. Ammer, G., Leonhardt, A., Bahl, A., Dickson, B.J., and Borst, A. (2015). Functional specialization of neural input elements to the *Drosophila* ON motion detector. *Curr. Biol.* *25*, 2247–2253.
78. Meier, M., Serbe, E., Maisak, M.S., Haag, J., Dickson, B.J., and Borst, A. (2014). Neural circuit components of the *Drosophila* OFF motion vision pathway. *Curr. Biol.* *24*, 385–392.
79. Lucas, B.D., and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. *Proceedings of the International Joint Conference on Artificial Intelligence* *2*, 674–679.
80. Plett, J., Bahl, A., Buss, M., Kühnlenz, K., and Borst, A. (2012). Bio-inspired visual ego-rotation sensor for MAVs. *Biol. Cybern.* *106*, 51–63.
81. Bagheri, Z.M., Cazzolato, B.S., Grainger, S., O'Carroll, D.C., and Wiederman, S.D. (2017). An autonomous robot inspired by insect neurophysiology pursues moving features in natural environments. *J. Neural Eng.* *14*, 046030.
82. Olsen, S.R., Bhandawat, V., and Wilson, R.I. (2010). Divisive normalization in olfactory population codes. *Neuron* *66*, 287–299.
83. Rabinowitz, N.C., Willmore, B.D.B., Schnupp, J.W.H., and King, A.J. (2011). Contrast gain control in auditory cortex. *Neuron* *70*, 1178–1191.
84. Baccus, S.A., and Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron* *36*, 909–919.
85. Busse, L., Wade, A.R., and Carandini, M. (2009). Representation of concurrent stimuli by population activity in visual cortex. *Neuron* *64*, 931–942.

86. Carandini, M., Heeger, D.J., and Movshon, J.A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* *17*, 8621–8644.
87. Tuthill, J.C., Nern, A., Holtz, S.L., Rubin, G.M., and Reiser, M.B. (2013). Contributions of the 12 neuron classes in the fly lamina to motion vision. *Neuron* *79*, 128–140.
88. Pologruto, T.A., Sabatini, B.L., and Svoboda, K. (2003). ScanImage: flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* *2*, 13. <https://doi.org/10.1186/1475-925X-2-13>.
89. Meyer, H.G., Schwegmann, A., Lindemann, J.P., and Egelhaaf, M. (2014). Panoramic high dynamic range images in diverse environments. <https://doi.org/10.4119/unibi/2689637>.
90. Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* *118*, 401–415.
91. Kingma, D.P., and Ba, J. (2014). Adam: A method for stochastic optimization. arXiv, arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>.
92. Richter, F.G., Fendl, S., Haag, J., Drews, M.S., and Borst, A. (2018). Glutamate signaling in the fly visual system. *iScience* *7*, 85–95.
93. Jia, H., Rochefort, N.L., Chen, X., and Konnerth, A. (2011). In vivo two-photon imaging of sensory-evoked dendritic calcium signals in cortical neurons. *Nat. Protoc.* *6*, 28–35.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Experimental data (behavior, electrophysiology, and calcium imaging)	This study	https://github.com/borstlab/normalization_paper
Experimental Models: Organisms/Strains		
<i>D. melanogaster</i> : WT: Canton S	N/A	N/A
<i>D. melanogaster</i> : L1-AD: w ¹¹¹⁸ ; VT027316-AD; +	Courtesy of A. Nern / Janelia Research Campus	N/A
<i>D. melanogaster</i> : L1-DBD: w ¹¹¹⁸ ; +; R40F12-DBD	Courtesy of A. Nern / Janelia Research Campus	RRID: BDSC_69935
<i>D. melanogaster</i> : L2-AD: w ¹¹¹⁸ ; R53G02-AD; +	[87]	RRID: BDSC_68990
<i>D. melanogaster</i> : L2-DBD: w ¹¹¹⁸ ; +; R29G11-DBD	[87]	RRID: BDSC_70173
<i>D. melanogaster</i> : L3-AD: w ¹¹¹⁸ ; R59A05-AD; +	[87]	RRID: BDSC_70751
<i>D. melanogaster</i> : L3-DBD: w ¹¹¹⁸ ; +; R75H07-DBD	[87]	RRID: BDSC_69459
<i>D. melanogaster</i> : L4-AD: w ¹¹¹⁸ ; R20A03-AD; +	[87]	RRID: BDSC_68957
<i>D. melanogaster</i> : L4-DBD: w ¹¹¹⁸ ; +; R31C06-DBD	[87]	RRID: BDSC_68978
<i>D. melanogaster</i> : L5-AD: w ¹¹¹⁸ ; R21A05-AD; +	[87]	RRID: BDSC_70588
<i>D. melanogaster</i> : L5-DBD: w ¹¹¹⁸ ; +; R31H09-DBD	[87]	RRID: BDSC_68980
<i>D. melanogaster</i> : Mi1-AD: w ¹¹¹⁸ ; R19F01-AD; +	[12]	RRID: BDSC_68955
<i>D. melanogaster</i> : Mi1-DBD: w ¹¹¹⁸ ; +; R71D01-DBD	[12]	RRID: BDSC_69066
<i>D. melanogaster</i> : Tm3-AD: w ¹¹¹⁸ ; R13E12-AD; +	[12]	RRID: BDSC_68830
<i>D. melanogaster</i> : Tm3-DBD: w ¹¹¹⁸ ; +; R59C10-DBD	[12]	RRID: BDSC_69153
<i>D. melanogaster</i> : Mi4-AD: w ¹¹¹⁸ ; R48A07-AD; +	[12]	RRID: BDSC_71070
<i>D. melanogaster</i> : Mi4-DBD: w ¹¹¹⁸ ; +; R13F11-DBD	[12]	RRID: BDSC_69722
<i>D. melanogaster</i> : Mi9-AD: w ¹¹¹⁸ ; R48A07-AD; +	[12]	RRID: BDSC_71070
<i>D. melanogaster</i> : Mi9-DBD: w ¹¹¹⁸ ; +; VT046779-DBD	[12]	RRID: BDSC_74714
<i>D. melanogaster</i> : Tm1-AD: w ¹¹¹⁸ ; R41G07-AD; +	[68]	RRID: BDSC_71049
<i>D. melanogaster</i> : Tm1-DBD: w ¹¹¹⁸ ; +; R74G01-DBD	[68]	RRID: BDSC_69767
<i>D. melanogaster</i> : Tm2: w ¹¹¹⁸ ; +; VT012282	[11]	N/A
<i>D. melanogaster</i> : Tm2split-AD: w ¹¹¹⁸ ; R28D05-AD; +	[68]	RRID: BDSC_68974
<i>D. melanogaster</i> : Tm2split-DBD: w ¹¹¹⁸ ; +; R82F12-DBD	[68]	RRID: BDSC_69250
<i>D. melanogaster</i> : Tm4: w ¹¹¹⁸ ; +; R35H01	[11]	RRID: BDSC_49922
<i>D. melanogaster</i> : Tm9: w ¹¹¹⁸ ; +; VT065303	[11]	N/A
<i>D. melanogaster</i> : T4-AD: w ¹¹¹⁸ ; VT016255-AD; +	Vienna Drosophila Resource Center	N/A
<i>D. melanogaster</i> : T4-DBD: w ¹¹¹⁸ ; +; VT012314-DBD	Vienna Drosophila Resource Center	N/A
<i>D. melanogaster</i> : T5-AD: w ¹¹¹⁸ ; VT013975-AD; +	Vienna Drosophila Resource Center	N/A
<i>D. melanogaster</i> : T5-DBD: w ¹¹¹⁸ ; +; R42H07-DBD	Bloomington Drosophila Stock Center	RRID: BDSC_69609
<i>D. melanogaster</i> : T4/T5-AD: w ¹¹¹⁸ ; R59E08-AD; +	[44]	RRID: BDSC_71101
<i>D. melanogaster</i> : T4/T5-DBD: w ¹¹¹⁸ ; +; R42F06-DBD	[44]	RRID: BDSC_69285
<i>D. melanogaster</i> : w+; P{20XUAS-IVS-GCaMP6f}attP40; +	Bloomington Drosophila Stock Center	RRID: BDSC_42747
<i>D. melanogaster</i> : w+; +; PBac{20XUAS-IVS-GCaMP6f}VK00005	Bloomington Drosophila Stock Center	RRID: BDSC_52869
<i>D. melanogaster</i> : UAS-TNT: +; UAS-TNT; +	[52]	N/A
<i>D. melanogaster</i> : UAS-TNTin: +; UAS-IMPTNT-Q; +	[52]	N/A

(Continued on next page)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Custom-written software in Python	This study	https://github.com/borstlab/normalization_paper
ScanImage 3.8	[88]	http://scanimage.vidriotechnologies.com/display/SlH/ScanImage+Home
Other		
Natural images for experiments and modeling	[43]	N/A
Natural images for modeling	[89]	https://doi.org/10.4119/unibi/2689637

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Aljoscha Leonhardt (leonhardt@neuro.mpg.de). This study did not generate new unique reagents.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Drosophila melanogaster were kept on a 12 h light/12 h dark cycle at 25°C and 60% humidity on standard cornmeal-agar medium. Genetic expression of effectors was targeted through the Gal4-UAS system [90]. Resulting genotypes and their abbreviations are listed in Table S2.

Unless stated otherwise, locomotion and tangential cell responses were recorded in wild-type Canton S flies 1 to 5 days after eclosion (Figures 1 and 2). We used the genetically encoded calcium indicator GCaMP6f [45] to determine the functional properties of individual cell types (Figures 3, 4, and 5). Throughout silencing experiments (Figure 5; Figure S2), we expressed tetanus toxin light chain (TNT) or an inactive version (TNTin) in the cell type of interest [52]. For calcium imaging experiments involving silencing (Figure 5), one day old flies were collected and put on 29°C for 3 days to boost expression of TNT or TNTin.

METHOD DETAILS

Natural image sets

For electrophysiology, behavioral, and modeling experiments, we used images from a published set of 20 natural panoramic scenes [43] termed dataset A. All images were independently processed as follows: We averaged across color channels and downsampled the scene to a resolution of 1,600 × 320 pixels (covering 360° sampled at 0.225 pixels per degree along the azimuth) using linear interpolation. To be able to render 12 bit images on conventional screens with 8 bits of dynamic range, we first performed standard gamma correction by raising raw pixel values to a power of 0.45 and then clipped the top percent of pixel intensities. The resulting image was scaled to fill the range between 0 and 255.

For optomotor experiments (Figure 1), we selected a subset of 8 images that covered different types of terrain. From this set, we again selected a subset of 6 images to determine tangential cell responses. We used all 20 images to build the convolutional network (Figure 6), randomly assigning 15 scenes to the training and 5 to the test set. Finally, we validated the trained convolutional model with images from an independent panoramic scene collection [89] consisting of 421 images (Figure S6G). These scenes were kept at their native resolution of 927 × 251 pixels (corresponding to an azimuthal sampling rate of 0.39 pixels per degree) and processed as above, yielding dataset B. We then generated two test sets: One had gamma correction applied to limit the images' bit depth ("low dynamic range" or LDR) and the other one was left at 12 bit depth to produce a dataset with high dynamic range (HDR).

Behavioral experiments

Experiments on the treadmill setup were conducted as described before [20, 31, 44]. Briefly, we tethered flies to a thin metal rod and placed them on air-cushioned polyurethane balls whose movement was tracked at 4 kHz, allowing for direct readout of rotational motion along all three axes. Temperature within the vicinity of the fly was 25°C at the start of each experiment. Using a closed-loop thermoregulation system, we linearly increased it to 34°C within 15 min to encourage locomotion.

For visual stimulation, we used three identically calibrated computer screens that were placed in a rectangle surrounding the fly. To simulate a cylindrical display, all stimuli were rendered onto a virtual cylinder and distorted accordingly before projection onto screens. Our setup covered approximately 270° in azimuth and 120° in elevation of the visual field. All stimuli were displayed at 144 Hz and at a spatial resolution greatly exceeding that of the fly eye. Screens had a maximum luminance of approximately 100 cd m⁻² and a luminance depth of 8 bit; for all descriptions below, we assume pixel brightness to range from 0 to a maximum of 1. Patterns were generated in real-time and programmed in Python 2.7 using the game engine Panda3D.

We measured velocity tuning curves (Figure 1) for 8 distinct natural images at 6 logarithmically spaced velocities ranging from 5 to $1,280^{\circ}\text{s}^{-1}$. Initial image phase was randomized on each trial. Scenes were displayed at their native gamma-corrected mean luminance and contrast (see above). On each trial, images stood still for 1.5 s, then were rotated at the chosen velocity for 0.5 s, and remained fixed for another 1.5 s.

The optomotor contrast stimulus separated the visual field into two areas (see Figure 2A; Figure S2). For the so-called background, we tiled the visual field with pixels of size $5^{\circ} \times 5^{\circ}$. At each pixel location we drew a temporal frequency f from a normal distribution ($\mu = 0$ Hz, $\sigma = 1$ Hz) and a starting phase λ from a uniform distribution covering 0 to 360° . Instantaneous luminance of each pixel i was then determined by a random sinusoid of the form

$$I_i(t) = 0.5 + 0.5 c_{bg} g(\sin(2\pi f_i t + \lambda_i))$$

where the experimental parameter c_{bg} runs from 0 to 100% and controls the effective contrast of the background. To increase average contrast in the visual field, we applied the compressive transform

$$g(x) = x \sqrt{\frac{1 + \alpha^2}{1 + \alpha^2 x^2}}$$

where $\alpha = 5$ determined the degree of curve flattening. Using this method, we generated stochastic and dynamic visual input at a controllable contrast level without introducing coherent motion (see Figure 2B).

The so-called foreground delivered a coherent motion stimulus driving the optomotor response. It consisted of two vertical stripes that were placed at plus and minus 90° from the frontal axis of the fly, each spanning 20° in azimuth and the full screen elevation. We again tiled each stripe with pixels covering an area of approximately $5^{\circ} \times 5^{\circ}$. For each pixel i , luminance was fixed over time and determined by

$$I_i(t) = 0.5 + 0.5 c_{fg} g(\sin(\lambda_i))$$

where the experimental parameter c_{fg} controls the effective motion contrast and λ was independently drawn from a uniform distribution covering 0 to 360° . The pixelated noise pattern smoothly wrapped around the azimuthal borders when moving. Note that for all instantiations of the stimulus, mean luminance across the visual field was 0.5. We verified that at typical scales of visual processing in *Drosophila* (approximated as a Gaussian filter with FWHM = 25° that covers a majority of the receptive fields of visual neurons; see [9]), variation in average luminance around this mean was small (Figure S2L).

For the basic contrast tuning experiment (Figures 2A–2E; see Video S1), we exhaustively measured combinations of logarithmically spaced values for c_{fg} (1.6, 3.1, 6.3, 12.5, 25, 50, and 100%) and c_{bg} (0, 25, 50, and 100%). At the beginning of each trial we simultaneously presented the dynamic background and the static foreground pattern. Between 1.5 and 2.0 s following stimulus onset, the foreground pattern moved at a fixed velocity of 50°s^{-1} . For oscillation experiments (Figures 2F–2J), the motion period was extended to 6 s. While the foreground pattern was moving, we sinusoidally modulated the contrast of either fore- or background between 0 and 100% around a mean value of 50% and at the specified temporal frequency (see Figure 2F; Figures S2A, S2C, and S2E; Video S2). When mapping the spatial extent of the contrast-induced modulation, we set the modulation frequency to 1 Hz and restricted the background pattern to two stripes of 10° width flanking each foreground pattern (see Figure S2A). The distance parameter (15, 17.5, 20, 22.5, 25, 27.5, 30, 35, or 40°) determined the separation between centers of foreground and background. In this experiment, we additionally measured a zero-contrast background condition to obtain an appropriate modulation baseline. Here, the motion stimulus had a contrast of 25% and luminance in the rest of the field was set to a uniform 0.5. Example traces in Figure 2G are taken from this spatial experiment (for distance 15° or no background). For the temporal experiments, we measured oscillation frequencies of 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, 8, and 10 Hz (Figure 2J). Background contrast was zero when measuring foreground tuning; for background tuning, foreground contrast was set to 25%.

All stimulus patterns were displayed twice throughout optomotor experiments, once in clockwise and once in counterclockwise direction of motion. We recorded multiple trials to obtain robust turning responses for each fly (15 trials for natural image stimuli, 20 for contrast tuning, 25 for oscillation stimuli). Presentation order was shuffled across conditions within any trial to mitigate adaptation effects. Individual experiments lasted between 60 and 120 min.

Electrophysiology

Our patch-clamp recordings from tangential cells followed established protocols [11]. Cell bodies of horizontal system (HS) units were targeted visually through a microscope. We confirmed their preferred direction by stimulation with oriented moving sine wave gratings before each experiment.

Visual stimulation was delivered using a cylindrical projector-based arena as previously described [9]. Briefly, the screen of the arena covered a viewing angle of the fly of 180° in azimuth and 105° in elevation. Stimuli were generated at a framerate of 180 Hz using green light spanning approximately 500 nm to 600 nm in wavelength. The maximum luminance this arena achieved was 276 ± 48 cd m^{-2} (mean \pm SD across devices). All visual stimuli were rendered using custom software written in Python 2.7 and the Panda3D framework. Membrane potential was recorded using custom software written in MATLAB (MathWorks, MA).

We measured tuning curves for 6 distinct natural image panoramas at 9 logarithmically spaced velocities ranging from 2.5 to 640°s^{-1} (Figure 1). On each presentation, the scene was displayed at a fixed phase, stayed still for 1 s, and then rotated horizontally for 3 s at the chosen constant velocity. Image movement was always in the preferred direction of the HS unit. We showed images at

their native gamma-corrected mean luminance and contrast (see above). Each condition was repeated 5 times. Conditions and trials were randomly interleaved to exclude adaptation effects along any stimulus dimension.

Calcium imaging

Calcium imaging experiments were performed using custom-built two-photon microscopes as described before [9]. The imaging acquisition rate was 11.8 Hz for all experiments, or 23.7 Hz for the experiment in Figure 4D, with imaging resolutions ranging from 32×32 to 64×128 pixels. Image acquisition was controlled using the ScanImage software (version 3.8) [88]. We prepared flies as previously described [9, 14]. Briefly, *Drosophila* were anesthetized on ice and glued onto an acrylic glass holder with the back of their head exposed to a perfusion chamber filled with Ringer's solution. Then the cuticula was surgically opened to allow optical access.

Stimuli were presented using the same projector system as in electrophysiological experiments, with additional long-pass filters (cut-off wavelength of 550 nm) in front of the projectors to spectrally separate visual stimulation from GCaMP fluorescence signals.

To identify receptive field (RF) positions of individual neurons, white noise stimuli of 3 min length were used (except for T4 and T5 cells, see below). The stimuli were pre-rendered at 60 Hz and generated as previously described [9]. Briefly, the spatial resolution of all white noise stimuli was 2.8° of visual angle corresponding to 64 pixels across the 180° screen. For all lamina cells, the same stimulus was used in order to provide a systematic description of their spatiotemporal filtering properties (Figure S3). This stimulus had a Gaussian autocorrelation with a standard deviation of approximately 45 ms in time and a contrast of 25% around a mean intensity value of 50 on an 8 bit grayscale. For some medulla cell types, variants of this stimulus with higher contrast or longer time constants were used if necessary to reliably locate their RFs on the arena. Specifically, we mapped RFs for Tm4, Mi4, Mi9 and Tm9 with a binary stimulus at 100% contrast and a temporal cut-off frequency of 1 Hz. For Mi9, we chose a 1D version of this stimulus, consisting of horizontal (1.5 min) and vertical bars (1.5 min) instead of pixels.

For T4 and T5, we relied on a novel stochastic motion noise stimulus to determine RF coordinates. First, we determined the preferred direction of an ROI using drifting gratings. Then we displayed a stimulus consisting of 20 randomly distributed 15° wide circular windows. Inside of each window, a 30° wavelength sine grating drifted at 30°s^{-1} in the preferred direction (Figure S3X). The positions of these 20 windows were changed and randomly chosen every second over 4 min. Reverse correlation of T4 and T5 responses with the area covered by those windows at a given time point yielded motion-sensitive RFs which were fit with a Gaussian to determine center coordinates (Figure S3Y). These were verified by presenting 25° windows containing full contrast drifting gratings at the estimated RF center and 6 hexagonally distributed positions around the center. Cells responded only to the grating in the RF center (Figure S3Z).

For the experiments shown in Figure 3, a 25° circular window around the RF center of a cell defined the foreground whereas the rest of the screen was defined as background. Before stimulus presentation, we verified that RF centers were sufficiently distant from the border of the screen to allow full display of the foreground. A drifting sine grating with 30° wavelength and a velocity of 30°s^{-1} was shown, starting with medium gray at the center of the RF and moving for 4 s after stimulus onset (see Video S3). The contrast of the grating was varied independently between background and foreground. A stimulus matrix of 7 foreground contrasts (1.6, 4, 8, 16, 32, 64 and 100%) and 6 background contrasts (0, 8, 16, 32, 64 and 100%) at a constant mean luminance of 0.5 was presented.

For the experiments shown in Figures 4A–4C, the foreground contrast was chosen depending on the cell type as the point where the suppression elicited by 100% background contrast (as measured in Figure 3) would be greatest. This was 16% for Mi1, 32% for Tm1, 100% for Tm2 and 64% for Tm3. The background had 100% contrast and 30° wavelength. We varied either its direction, its velocity (0, 0.25, 0.5, 1, 2, 4, 8, 16, 32 or 64°s^{-1}), or restricted its presentation to an annulus with changing outer diameter. A reference condition with 0% background contrast was added to the stimulus protocol.

For the contrast-step stimulus experiments shown in Figure 4D (see Video S4), the background grating had 30° spatial wavelength, drifted with 90°s^{-1} after motion onset and its initial phase was randomized. For Tm2 it had full contrast, for Tm3 44% contrast. The 25° foreground window was 50% gray and we placed a 5° wide dot in the center. For Tm3, the dot was initially black and set to white for a duration of 50 ms at a given time interval after motion onset of the background grating. For Tm2, the dot was initially white and then set to black. The time interval was varied in steps of 50 ms from -250 ms to 500 ms and then in steps of 100 ms. Negative values indicate that the surround grating started to move after the dot changed its intensity. Additional time intervals were -500 ms and -1 s. The block experiments in Figure 5 were performed with the same frequency tuning stimuli as before (Figure 4B). For the contrast tunings, the same stimuli as in Figure 3 were used but with background contrast of either 0 or 100% only.

All stimuli were repeated three times in randomized condition order to prevent adaptation to any stimulus features.

Modeling

Natural motion stimuli

To evaluate the performance of our models under naturalistic conditions, we generated a synthetic set of motion sequences that closely mimicked the experimental stimuli described above. For each sequence we translated 360° images at a fixed horizontal velocity through a virtual window spanning 100° in azimuth. Given their panoramic nature, scenes wrapped around seamlessly at each border. Movies were generated at a time resolution of 100 Hz. To reduce jitter for small velocities, we linearly interpolated non-integer pixel shifts. Fly eye optics were simulated ahead of time. We blurred each frame with a Gaussian filter (full width at half-maximum of 4°) to approximate the acceptance angle of each photoreceptor [26] and then sampled individual signals from a rectangular grid with isotropic spacing of 4° (yielding 23×17 receptor signals per frame for dataset A and 23×23 for dataset B, as described above).

For the comparison in [Figure 1](#), we modeled the exact stimulus parameters of the electrophysiological experiment including an approximation of the image's starting phase on the arena. We generated sequences for our convolutional detector models ([Figure 6](#)) as follows: The set of 20 panoramic images was randomly split into a training group consisting of 15 scenes and a test group consisting of 5 scenes. For each sequence, a random image was drawn from the appropriate set. The stimulus lasted 5 s. Between 1 and 4 s, scene velocity stepped from zero to a fixed value drawn from a Gaussian distribution with $SD = 100^\circ s^{-1}$. The initial window phase followed a uniform distribution spanning 360° . To further augment the dataset, we flipped the underlying image along the horizontal and vertical axes with a probability of 50%. We generated 8,192 such sequences for the training set and 512 for the test set.

Experimental stimuli

For all modeling experiments in [Figure S5](#), we replicated the experimental protocols described above as precisely as feasible. All stimuli were projected onto a field of view that spanned 120° in azimuth and 90° in elevation at a spatial resolution of 1° for calcium imaging experiments and 0.5° for behavioral experiments. Frames were then blurred and sampled as described for natural image stimuli. Brightness values for all stimuli ran from 0 to 1 and we fixed the mean level for contrast stimuli at 0.5. For calcium imaging stimuli, we always placed the foreground disk at the center of the field of view. Patterns were rendered and processed at 100 Hz.

Tuning curves for the basic contrast experiment ([Figures S5B–S5D](#)), the frequency experiment ([Figure S5F](#)), and the background diameter experiment ([Figure S5G](#)) were estimated from a single trial per parameter setting. For the background orientation experiment ([Figure S5E](#)) and the step interval experiment ([Figure S5H](#)) we averaged 100 trials with randomized background pattern phases to approximate the experimental phase stochasticity that results from individual cell receptive fields being located in different parts of the visual field. We averaged 200 trials for the behavioral stimuli ([Figure S5K](#)) to account for the intrinsic stochasticity of the stimulus and to generate reliable model responses. Throughout [Figure S5](#), we calculated point estimates for all tuning curves exactly as described for the behavioral and calcium data.

Tuning curve normalization model

The analytical model for divisive normalization ([Figures 3V–3X](#)) resembles previous formulations in the literature [[48](#), [50](#), [86](#)]. The steady-state response R of a neuron is given by

$$R(c_{fg}, c_{bg}) = \frac{L_{fg}c_{fg}^p + L_{bg}c_{bg}^p}{c_{50}^p + c_{fg}^p + S^p}$$

where c_{fg} and c_{bg} are foreground and background contrast and L_{fg} and L_{bg} are weight factors defining the respective amount of linear contribution of foreground and background to the response. The semi-saturation constant c_{50} determines the contrast at which the cell responds with 50% strength and the parameter p defines the steepness of the saturation curve.

The normalization term

$$S = w_{pool} \cdot c_{bg}^q$$

gives the amount of divisive surround suppression which is proportional to background contrast to a power of q , which accounts for possible non-linear scaling behavior, with a proportionality weight constant w_{pool} . In this model, the normalization index w_{pool}/c_{50} quantifies how much the sigmoidal tuning curve shifts to the right when c_{bg} is increased from 0 to 1 (full contrast), in relation to the semi-saturation constant. It thus describes the fold decrease in contrast sensitivity between no background contrast and full background contrast.

For evaluation of the normalization index ([Figure 3X](#)), this model was fit individually for each cell. Parameter fits to the average tuning curve per cell type are listed in [Table S1](#). Since tuning curves from individual cells are subject to measuring inaccuracies, we cross-validated fit quality. We optimized model parameters for the average tuning curve of 50% of all measured cells per type and evaluated variance explained for the other 50%. This was repeated 100 times with shuffled training and validation sets. For all cell types, cross-validated variance explained was more than 90% (see $R^2_{\text{DivisiveNorm}}$ in [Table S1](#)). When we repeated this procedure with a fully linear model

$$R(c_{fg}, c_{bg}) = L_{fg}c_{fg} + L_{bg}c_{bg}$$

variance explained dropped substantially for all units except L3 (see R^2_{linear} in [Table S1](#)).

This analysis was implemented using Python 2.7 and NumPy 1.11.3. Optimization of model parameters was performed using the L-BFGS-B algorithm in SciPy 0.19.0.

Data-driven detector model

The reference model in [Figure 1](#) was based on a standard implementation of the Reichardt-type correlational motion detector [[26](#)]. Briefly, all receptor signals of the two-dimensional input grid (see above) were filtered with a first-order high-pass ($\tau = 150$ ms). We then multiplied each local signal with the delayed horizontal neighbor (first-order low-pass, $\tau = 50$ ms). This was done twice in a mirror-symmetrical fashion and resulting output was subtracted. Finally, we summed across all local detectors to derive a model of tangential cell output. For the illustration in [Figure S1C](#), we simulated the receptor array at the full image resolution without blurring. These models were implemented in Python 3.6 using PyTorch 0.4.1.

We simulated time-resolved cell models for three basic response types: a purely linear low-pass unit (modeled after L3; [Figure S5B](#)), a strongly normalized band-pass unit (modeled after Mi1; [Figure S5C](#)), and a weakly normalized low-pass unit (modeled after Mi9; [Figure S5D](#)). We hand-tuned parameters based on our and previous work [[9](#)] to qualitatively match response properties

of the corresponding cell. Models were implemented as signal processing cascades (see Figure S5A). First, signals at each location in the field of view were filtered with a spatial difference of Gaussians kernel that had a central full-width at half-maximum (FWHM) of 6° and a FWHM of 20° in the surround. In accordance with results from receptive field mapping (Figure S3), the weight ratio between surround and center was 100% for low-pass units and 50% for the band-pass model. Full-field flashes would thus produce no activation in low-pass units. This was followed by first-order temporal filters: a single low-pass filter for low-pass units ($\tau = 80$ ms) or serial low- ($\tau = 50$ ms) and high-pass filters ($\tau = 150$ ms) for band-pass units. We then left the signal as is for ON cells or sign-inverted it for OFF cells and half-wave rectified the output by setting all negative values to zero.

For normalized cell models, we calculated local input P_i from the normalization field by pooling across rectified signals x_i with a Gaussian kernel (FWHM = 30°). Final output was then calculated using the divisive normalization equation

$$f(x_i) = \frac{x_i^p}{c_{50}^p + x_i^p + (w_{pool} P_i)^p}$$

where i indexes across points in space and time, c_{50} determines baseline sensitivity, exponent p regulates the static response non-linearity, and w_{pool} adjusts sensitivity to the normalization field signal. We manually tuned normalization parameters for the band-pass ($c_{50} = 0.012$, $p = 1.3$, $w_{pool} = 1.5$) and the low-pass cell ($c_{50} = 0.12$, $p = 1.1$, $w_{pool} = 3.0$) to match critical features of the empirical contrast tuning curves (Figures S5C and S5D).

To generate simulated T4 responses (Figures S5I and S5J), we multiplied the output of spatially adjacent low- and band-pass units. For the linear reference model we bypassed the final normalization step in both arms of the detector. We built the LPTC model (Figure S5K) as a spatial array of T4 and T5 cells covering the full field of view, analogously to the previously described two-quadrant detector [28]. For the T5 model, we used two OFF-sensitive input units with identical parameters as for ON cells. Output from syndirectionally tuned T4 and T5 motion detectors was summed and subtracted from a mirror-symmetric, oppositely tuned array to produce LPTC model output. The same model was used to simulate natural scene responses (Figures S5L–S5N). All models in Figure S5 were implemented using Python 3.6 and NumPy 1.15.

To quantify the robustness of velocity tuning for models and LPTCs (Figure S5N), we calculated per-velocity coefficients of variation as the ratio between response standard deviation across images and response mean across images. For neural data, we used cell-averaged mean potential to estimate these parameters.

Task-driven detector model

We implemented the trained detector model as a four-layer convolutional neural network consisting of linear input filters, a normalization stage, local multiplication, and linear spatial summation. In contrast to typical deep architectures used for object recognition, this network processed three-dimensional inputs spanning two dimensions of space as well as time.

First, receptor signals of shape $23 \times 17 \times 500$ or $23 \times 23 \times 500$ (azimuth, elevation, time), depending on the dataset, were processed in two independent convolutional channels. The convolutions were temporally causal and spatiotemporally separable. Each of the channels was composed of a $3 \times 3 \times 1$ spatial filter (covering 3 simulated receptors in azimuth and elevation) followed by a temporal filter of shape $1 \times 1 \times 30$ (corresponding to 300 ms at the chosen time resolution of 100 Hz). Convolutions had no bias parameter. In contrast to standard Reichardt detectors, each filter weight was allowed to vary freely during optimization.

Second, we passed local output signals x_i (where i indexes points in space and time) through one of three types of local normalization: a simple pass-through (termed “linear”)

$$f(x_i) = x_i$$

a static and contrast-independent compression stage (termed “static”)

$$f(x_i) = \tanh\left(\frac{x_i}{c}\right)$$

where the trained parameter c determines the sensitivity of the saturating function, or an adaptive saturation stage (termed “dynamic”)

$$f(x_i) = \tanh\left(\frac{x_i}{c + P_i}\right)$$

where c again determines the baseline sensitivity and P_i is the instantaneous output of a $11 \times 11 \times 1$ spatial filter (centered on the location of x_i and operating on full-wave rectified output signals $|x_i|$; see Figure 6A). This models the fast and spatially distributed normalization we observed during experiments. We chose the hyperbolic tangent because it generalizes to positive and negative input values, the transformation closely resembles the normalization model described above, and it is more commonly used in the field of deep learning. Spatiotemporal filters were optimized independently for each of the two channels while the sensitivity parameter c was shared.

Third, we then combined signals from both channels in a EMD-type scheme where adjacent signals were multiplied and output from two mirror-symmetric pairs was subtracted. This stage was parameter-free. Finally, resulting signals were summed across space and multiplied by a trained scalar amplification factor to generate the final time-resolved output of the model. The base model without normalization had 79 trainable parameters; static normalization added one parameter and dynamic normalization another 242.

We trained each model architecture to estimate the true velocity of translation stimuli using automatic differentiation, backpropagation, and stochastic gradient descent. The loss function we applied was the mean squared error (MSE) between model output and current velocity of the scene. Weights were updated using the Adam optimizer [91], with parameters set to standard values ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$). Models were trained over 800 epochs with a batch size of 128; no early stopping was used. We set the initial learning rate to 0.025 and divided it by a factor of 4 after 400, 500, and 600 steps. Input convolutional layers were initialized to random values drawn from a uniform distribution. For the pooling receptive field, we initialized each weight with 0.0001 and the sensitivity factor c with 1.0. Static sensitivity as well as pooling weights were constrained to be positive. In the dynamic normalization model, we applied a L_2 penalty of 400.0 to the spatial weights of the pooling stage. Hyperparameters were determined in preliminary experiments with an independent image set. We optimized each architecture 16 to 23 times with different random number generator seeds to assess reliability and did not select models post hoc.

We implemented all architectures in Python 3.6 using PyTorch 0.4.1 for automatic differentiation. Depending on model type, a single optimization run took between 6 and 14 hs on an NVIDIA Titan Xp GPU.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data evaluation for behavioral experiments

To ensure data quality, we excluded all flies whose average forward velocity during the experiment was below 0.25 cm s^{-1} and whose average turning tendency was either slowly drifting or far from 0° s^{-1} . Fewer than 20% of all experiments failed these criteria. Measurements of ball movement were downsampled via linear interpolation for further processing (to 50 Hz for natural image stimuli, Figure 1; 20 Hz for contrast tuning, Figure 2; 100 Hz for oscillation stimuli, Figure 2). Trials were averaged.

Responses for clockwise and counterclockwise motion were subtracted and divided by two to minimize residual deviations from straight forward walking. Traces for natural image and contrast tuning stimuli were filtered using a first-order low-pass with a time constant of 100 ms. For the contrast oscillation experiments, we evaluated modulation at the relevant carrier frequency by calculating the zero-padded Fourier Transform of the turning trace and averaging the amplitude spectrum in a window of width 0.2 Hz centered on the target frequency. These values were normalized per experiment such that the modulation peak after averaging was 100%. We applied a Savitzky-Golay filter (window length 11 samples, 5th order polynomial) before plotting traces from oscillation experiments; this did not affect the analysis.

All analysis for behavioral experiments was performed in custom-written software using Python 3.6, NumPy 1.15, and SciPy 1.1.

Data evaluation for electrophysiological experiments

Voltage data were digitized at 1,000 Hz. To account for slow drift in potential, we subtracted the average voltage in a 1 s window before stimulus onset from each trace per stimulus condition and trial. Signals were then low-pass filtered (8th order Chebyshev Type 1) and resampled at 100 Hz. Finally, we averaged cell responses across trials. Cells whose mean depolarization during full-contrast sine grating presentation in preferred direction remained below 5 mV were discarded before further analysis. All analysis for electrophysiological experiments was performed in custom-written software using Python 3.6, NumPy 1.15, and SciPy 1.1.

Data evaluation for calcium imaging experiments

Calcium imaging stacks were registered in order to correct for translational movement artifacts of brain tissue using custom-written software. Responses of individual neurons were extracted by manually selecting small regions of interest (ROI) encompassing individual anatomical structures. For T4 and T5 these corresponded to single or few axon terminals; for Mi and Tm cells, individual axon terminals could be identified clearly through visual inspection. For ON pathway medulla cells, signals were measured in layer 10 of the medulla, for OFF pathway medulla cells in layer 1 of the lobula. For lamina cells L1–5, signals were measured at axon terminals in corresponding layers 1–5 in the medulla. For T4 and T5, signals were recorded in the lobula plate.

To reconstruct RFs, calcium signals were mean subtracted and reverse-correlated with the stimulus as previously described [9]. 1D Gaussians were fit to horizontal and vertical cross-sections of spatial receptive fields to obtain precise RF coordinates. For lamina cells (Figure S3), all reconstructed RFs were peak-aligned and analyzed as previously [9]. For 1D projections of spatial RFs (Figures S3F–S3J), an average of 1D projections of 2D RFs along 3600 evenly distributed projection angles between 0° and 360° was calculated. This enhanced the visibility of the center-surround structure but neglected possible anisotropies in the spatial structure of RFs [49]. For impulse responses (Figures S3K–S3O) the temporal receptive field of the 9 center pixels was averaged; frequency responses (Figures S3P–S3T) are the Fourier-transformed impulse responses. Deconvolution (Figures S3U and S3V) was performed by dividing the frequency spectra with the frequency response of a 1st order low-pass filter with time-constant 350 ms as a proxy for calcium indicator dynamics [9, 92].

Relative fluorescence changes ($\Delta F/F$) from raw calcium traces were obtained by adapting an automatic baseline detection algorithm [93]. Briefly, raw data were first smoothed with a Gaussian window (full-width at half maximum, FWHM = 1 s). Then, minima within a 90 s long sliding window were extracted and the resulting trace smoothed with a Gaussian window (FWHM = 4 min). The result was used as a dynamic baseline F_0 and $\Delta F/F$ values were computed as $\Delta F/F = (F - F_0)/F_0$.

For further evaluation, only recordings with good signal-to-noise ratio (SNR) were taken. The criterion was that the standard deviation of the mean signal averaged over trials had to be at least 120% of the mean standard deviation over trials. This criterion filtered out cells with an inter-trial variance larger than the typical cell response (caused by movement artifacts or photobleaching).

In addition, the standard deviation of the mean signal had to be larger than 25% $\Delta F/F$. On average, 90% of all cells measured passed these criteria with slight variations due to different levels of GCaMP expression depending on the genotype.

For experiments with drifting gratings, the driving foreground contrast frequency was 1 Hz. For these experiments, we evaluated the amplitude of the 1 Hz component of the signal. This was achieved by computing the Fourier coefficient at that frequency, using the equation

$$F = \left| \frac{1}{T} \int_0^T dt s(t) e^{-2\pi i \cdot 1\text{Hz} \cdot t} \right|$$

where $s(t)$ denotes the signal and T the stimulation time. For experiments in Figure 4D, we evaluated the peak response of the calcium signal. For Figure S4, we additionally evaluated the average calcium signal (F_0) during stimulus presentation and normalized it to the maximum amplitude of the 1 Hz component (F_1).

Amplitudes were averaged over trials and normalized to the maximum, then averaged over cells and normalized to the maximum. For Figures 4 and 5, amplitudes were normalized to the response amplitude for the reference stimulus.

Statistical tests

Unless indicated otherwise, error bars show bootstrapped 68% confidence intervals around the mean (estimated as corresponding distribution percentiles after resampling the data 1,000 times). All statistical tests were two-tailed and performed at a 5% significance level. Normality of data distributions was assessed visually but not tested formally. Sample sizes are given in each figure legend and were not based on power analysis but predetermined in line with standards in the field. We did not blind experimenters to genotypes or conditions during data gathering and analysis.

DATA AND CODE AVAILABILITY

Code and experimental data are available on GitHub (https://github.com/borstlab/normalization_paper).

Current Biology, Volume 30

Supplemental Information

Dynamic Signal Compression

for Robust Motion Vision in Flies

Michael S. Drews, Aljoscha Leonhardt, Nadezhda Pirogova, Florian G. Richter, Anna Schuetzenberger, Lukas Braun, Etienne Serbe, and Alexander Borst

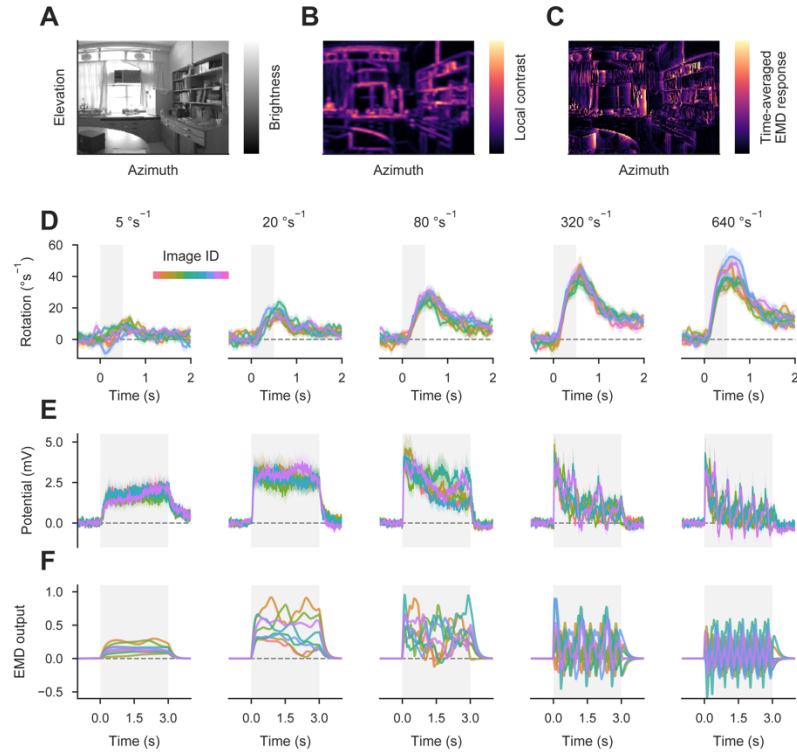


Figure S1 | Behavioral, neural, and model responses to natural scenes. Related to Figure 1.

(A) Natural image patch as seen through the field of view of model LPTC. (B) Estimate of local contrast in natural image patch. RMS contrast was estimated by filtering the image with a Gaussian ($\sigma = 0.5^\circ$), subtracting the filtered image from the original, squaring the mean-subtracted image, filtering it with a Gaussian ($\sigma = 0.5^\circ$), and taking the square root. (C) Spatially reconstructed output of simulated LPTC for same image patch as before, plotted as the square root of the time-averaged response. A horizontally motion-sensitive LPTC was constructed using the same parameters as in Figure 1F (STAR Methods) with the exception of more fine-grained sampling at exactly the image resolution. The depicted panorama was moved for 16 s at a velocity of 22.5°s^{-1} , resulting in a single complete rotation. Responses at each pixel location were then averaged across the full stimulus period. This demonstrates that the response of the EMD array depends strongly on squared local image contrast. (D) Turning responses for 8 images (indicated by trace color) and 5 velocities (indicated by panel title; $N=16$ wild-type flies; data as in Figure 1B). Gray shaded area indicates duration of motion stimulus. (E) Membrane potential for 6 images and 5 velocities ($N=11$ HS cells from 9 flies; data as in Figure 1D). (F) Output of model LPTC for same images and velocities as E (data as in Figure 1F). See Table S2.

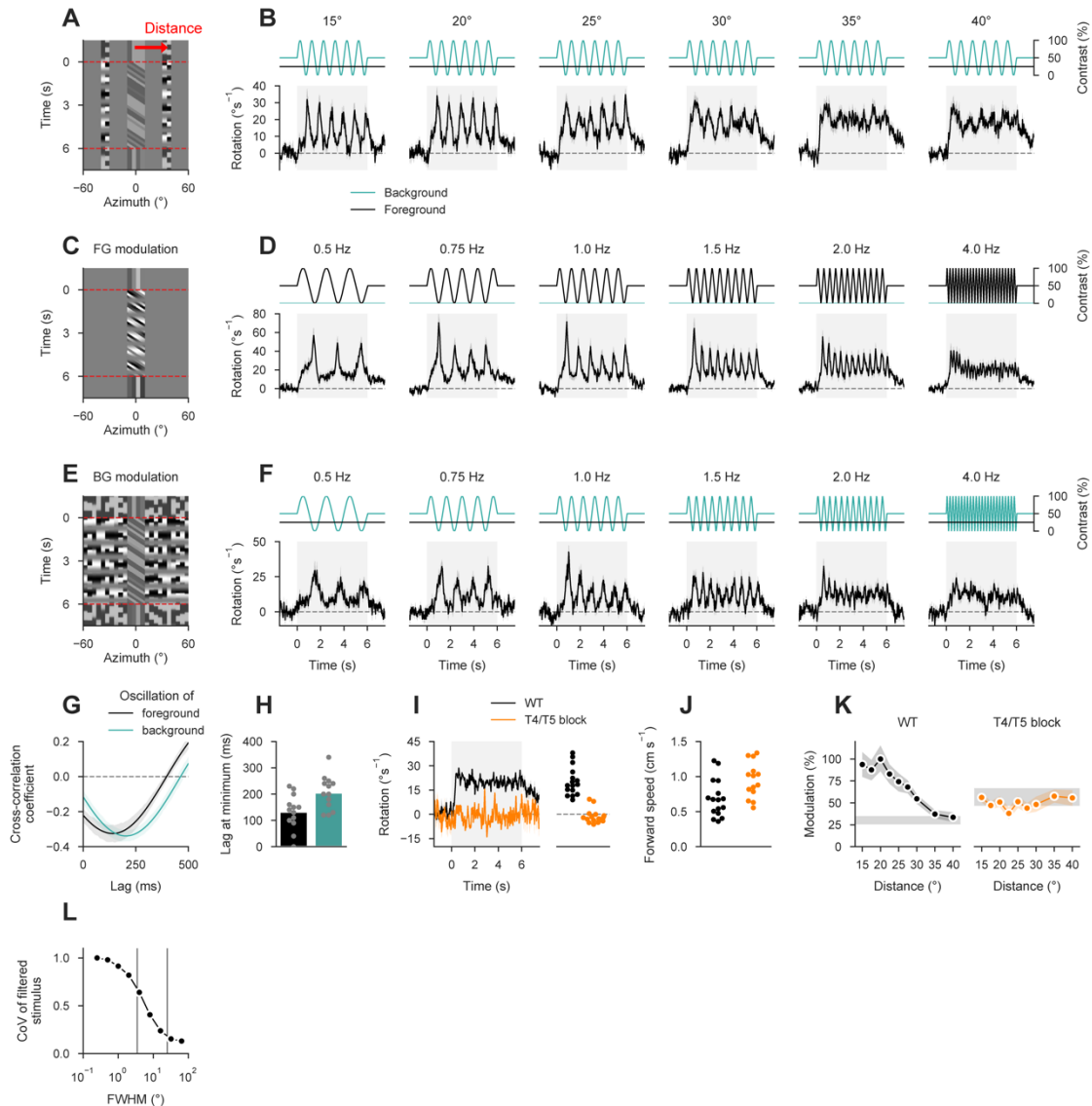


Figure S2 | Detailed behavioral responses to contrast stimuli. Related to Figure 2.

(A) Illustration of spatial oscillation experiment. Background was restricted to 10° wide stripes flanking the foreground motion stimulus at the center distance indicated by the red arrow. Dashed lines indicate period during which foreground pattern moved at 50°s^{-1} . This arrangement was repeated at plus and minus 90° from the frontal axis of the fly; 0° in this plot indicates the center of the foreground. (B) Contrast traces and turning responses for five distance conditions (indicated above each panel). Top, instantaneous contrast (25% in foreground, oscillating at 1 Hz in background). Bottom, turning response of the fly (N=16 wild-type flies). Modulation was reduced as spacing between foreground and background increased. (C) Illustration of temporal foreground modulation stimulus at 1 Hz frequency. (D) Contrast traces and turning responses for five foreground oscillation frequencies (N=13; background contrast was 0%). Modulation decreased as frequency increased. (E) Illustration of temporal background modulation stimulus at 1 Hz frequency. (F) Contrast traces and turning responses for five background oscillation frequencies (N=13; foreground contrast was 25%). Modulation again decreased with frequency. (G) Normalized cross-correlation between contrast oscillation and turning behavior for 1 Hz data from D and F. (H) Lag between stimulus oscillation and turning, evaluated as per-fly lag at first minimum within 500 ms for cross-correlations from G. (I) Left, comparison of turning responses between wild-type flies and flies in which T4/T5 cells were silenced using TNT (STAR Methods; N=16/14 for WT/block flies). Right, turning responses averaged between 0 and 6 s following motion onset. Syndirectional turning was abolished in T4/T5-silenced flies. (J) Average forward speed throughout full experiment. T4/T5 block flies did not exhibit

locomotion deficiencies. **(K)** Comparison of spatial oscillation tuning. T4/T5 block flies did not show modulation at the contrast oscillation frequency of 1 Hz and a generally increased level of baseline fluctuation. **(L)** Evaluation of luminance properties at different spatial scales for the behavioral stimulus. Normalized coefficient of variation across visual field was calculated after applying a Gaussian filter with different full-widths at half-maximum (FWHM). Gray lines indicate typical FWHM of ommatidium (left) and full receptive field of medulla cells (right). See Table S2.

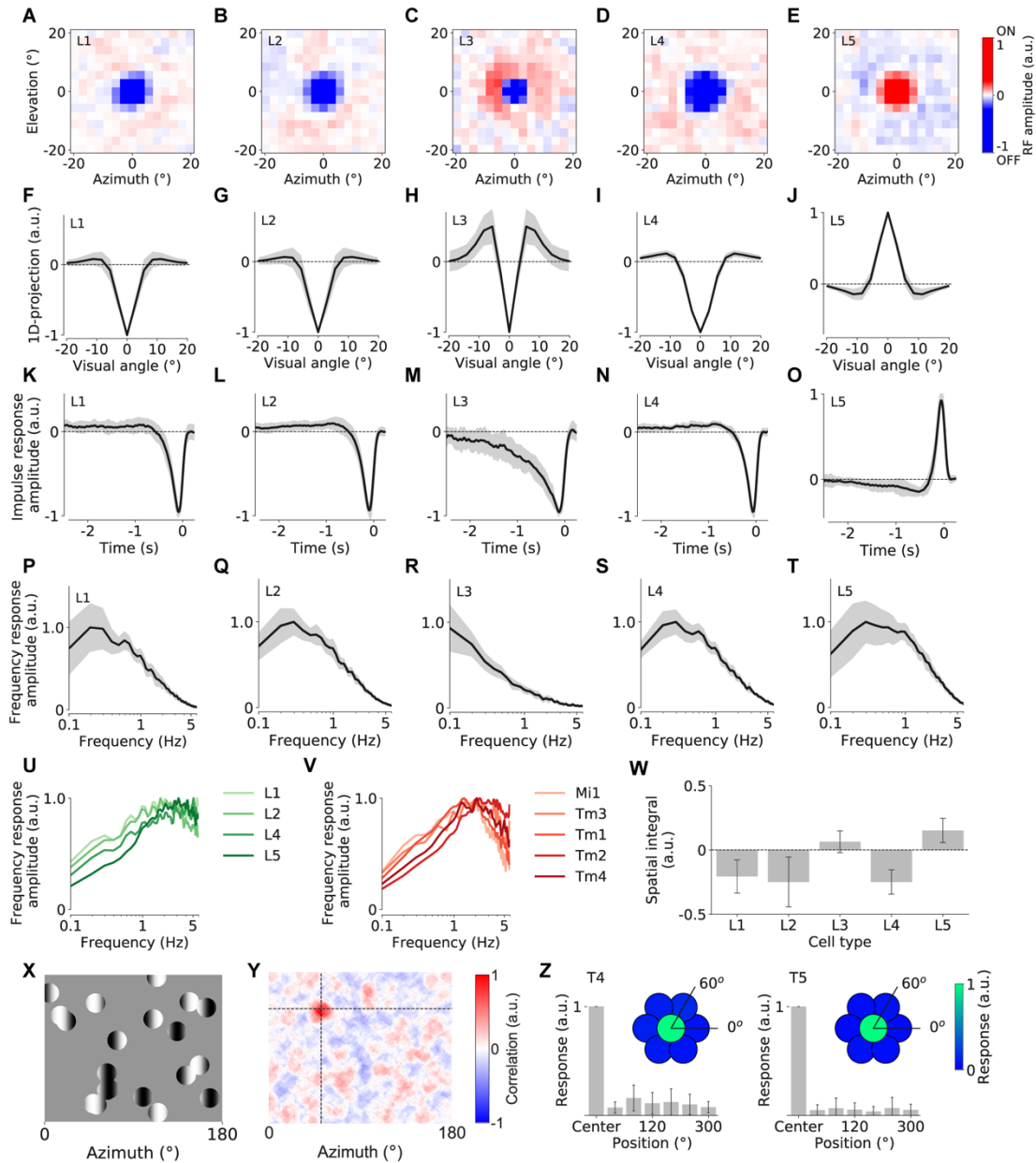


Figure S3 | Lamina and T4/T5 receptive field mapping. Related to Figure 3.

(A–E) Averaged 2D spatial receptive fields (RF) of L1–L5 from reverse correlation using white noise stimulation (L1: 21/7 cells/flyes, L2: 34/5, L3: 34/5, L4: 17/6, L5: 18/9). (F–J) 1D projection (averaged over all orientations) of the RFs in A–E. All cell types possessed linear RFs with antagonistic center-surround structure. (K–O) Temporal RFs measured in the center of the spatial RFs. (P–T) Frequency-space representations of temporal RFs. (U) Frequency representations of lamina transient cells (all lamina cells except for L3) after deconvolution with a putative linear GCaMP6f low-pass filter with time constant 350 ms as performed previously [S1]. (V) Deconvolved frequency responses of medulla bandpass filter cells (replotted from previous work [S1]). (W) Spatial integral of the 2D RFs in A–E. For L3, the strong antagonistic ON surround exactly counterbalanced the OFF-center contribution. (X) x-y plot of the stochastic motion noise stimulus used for localizing T4/T5 RFs. (Y) Example RF of a T4 cell from reverse correlation with the motion noise stimulus. (Z) Average responses of T4/T5 to 25° windowed drifting gratings probing different positions around the estimated RF center. This validated the RF coordinates obtained from the stochastic motion noise stimulus. All data are shown as mean ± s.d. See Table S2.

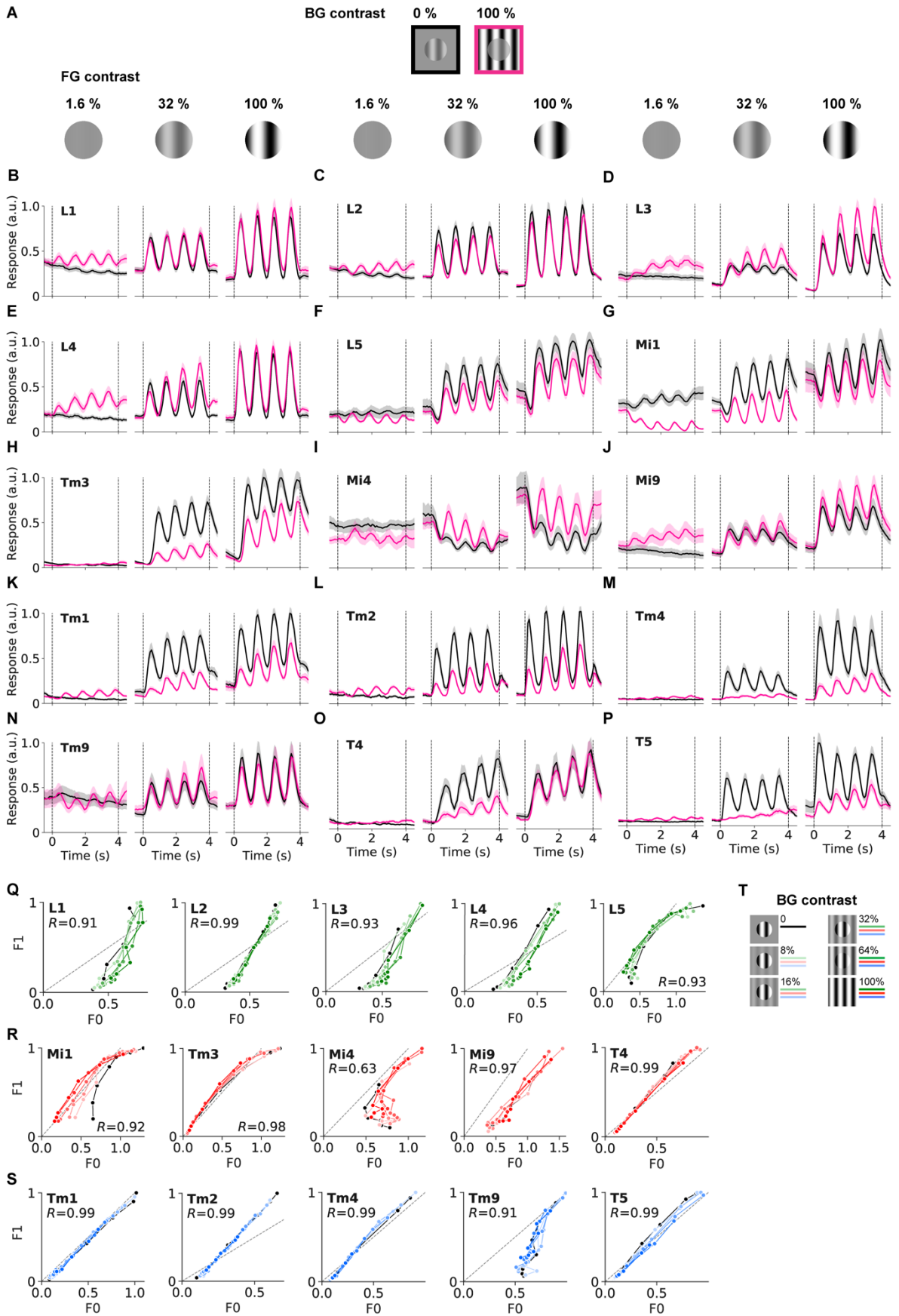


Figure S4 | Raw calcium responses for basic contrast stimuli. Related to Figure 3.

(A) Shown is only a subset of the data evaluated in Figure 3. Background contrast of 0 % is indicated by black lines, background contrast of 100 % is depicted in magenta. Responses are shown only for 3 out of 7 foreground contrasts. (B–P) Average calcium responses of all neurons to combinations of different foreground and background contrasts. (Q) Shown is a correlation analysis of the same dataset as in Figure 3 for the lamina cells L1–L5. On the y-axis is the F1-component of the calcium response (as evaluated in Figure 3) while the x-axis indicates the F0-component of the signal, i.e. the average calcium response during the stimulus. Data points corresponding to the same BG contrast are connected by lines and color-coded analogously to Figure 3. All data points are normalized to the maximum F1 response for each cell type. The gray dashed line marks the diagonal of the coordinate system. Correlation coefficient R is indicated in each panel. (R) Same as in Q but for ON-pathway medulla cells Mi1–Mi9 and for T4 cells. (S) Same as in Q but for OFF-pathway medulla cells Tm1–Tm9 and for T5 cells. (T) Color legend for panels Q–S. Darker color shade corresponds to higher background contrast, similarly to Figure 3. Zero background contrast condition is shown in black. See Table S2.

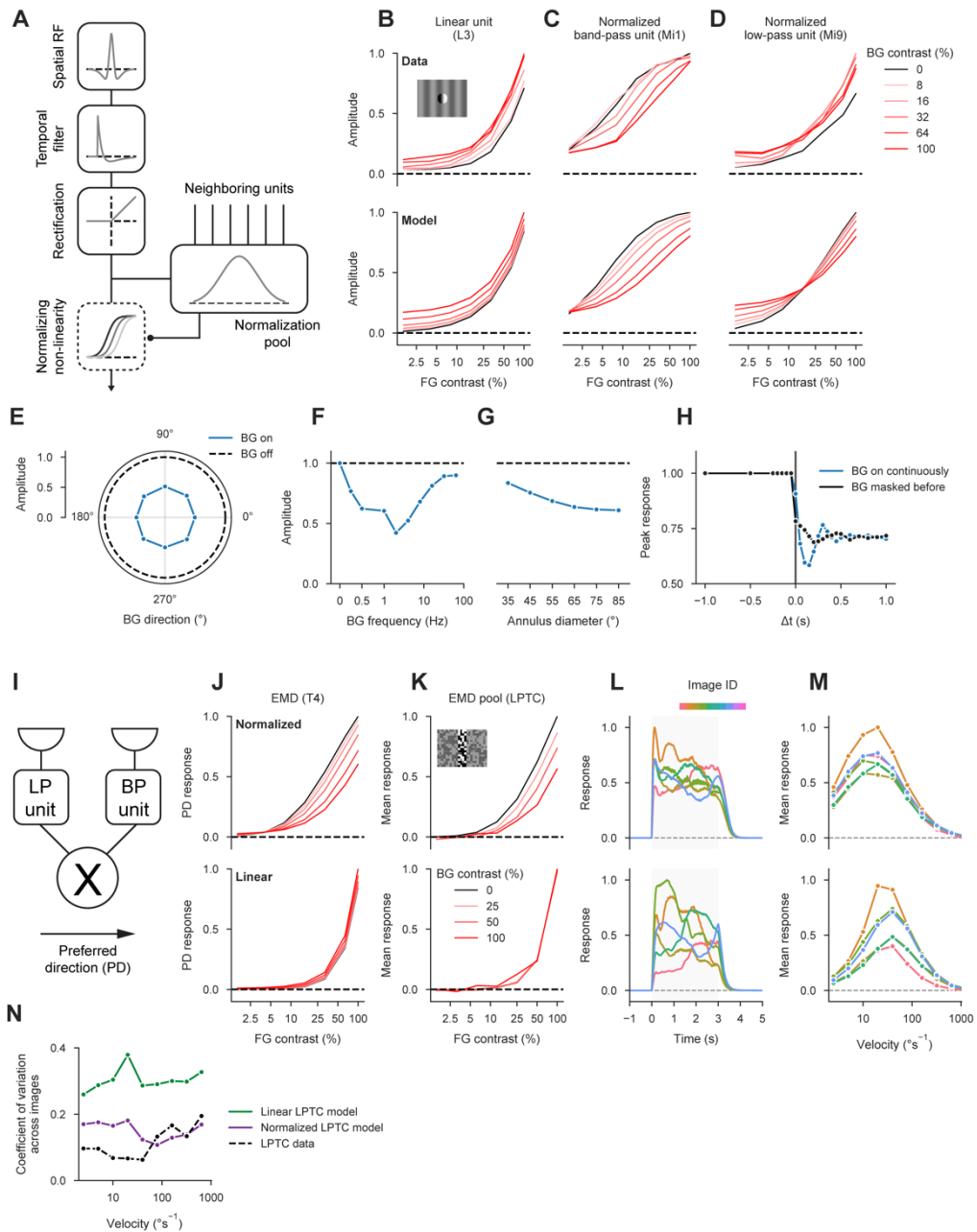


Figure S5 | Data-driven functional model of normalization circuit. Related to Figure 4.

(A) Illustration of signal cascade for data-driven cell model (STAR Methods). Filter elements are sketched for an ON band-pass cell with normalization. (B–D) Contrast tuning curves for three model cells, estimated using the same protocol as during calcium imaging (FG = foreground, BG = background). Top, empirical data for L3, Mi1, and Mi9 (see Figure 3). Inset depicts a single frame from stimulus centered on recorded cell with background contrast 25 % and foreground contrast 100 %. Bottom, tuning curves from models manually tuned to resemble their empirical counterparts (see STAR Methods for parameters). (E) Responses of normalized ON band-pass cell model to orientation tuning stimulus (see Figure 4A; dashed line marks reference stimulus without background). Stimuli and evaluation were exactly matched to the experiment. (F) Responses of the same model to background frequency tuning experiment (see Figure 4B; dashed line marks reference stimulus without background). (G) Responses of the same model to background size stimulus (see Figure 4C;

dashed line marks reference stimulus without background). **(H)** Responses of the same model to contrast-step protocol (see Figure 4D). **(I)** Illustration of T4 or T5 model. Signals from a strongly normalized band-pass and a weakly normalized low-pass unit covering adjacent areas of the visual field are multiplied, yielding a direction-selective signal. **(J–M)** Top, responses from motion detector models with normalization. Bottom, responses from motion detector models in which normalization was switched off for both input arms. **(J)** Foreground contrast tuning for simulated T4 cell (see Figure 3). **(K)** Responses to behavioral contrast stimulus for a LPTC model composed of T4 and T5 models (STAR Methods). **(L)** Responses to various natural scenes moving at 20°s^{-1} (modelled and evaluated as in Figure 1). **(M)** Velocity tuning curves for natural scenes (modelled and evaluated as in Figure 1). **(N)** Coefficient of variation across images for individual image velocities (derived from velocity tuning curves in M and Figure 1F; STAR Methods). A model including input normalization outperformed the linear model and approximated the variability of LPTC responses.

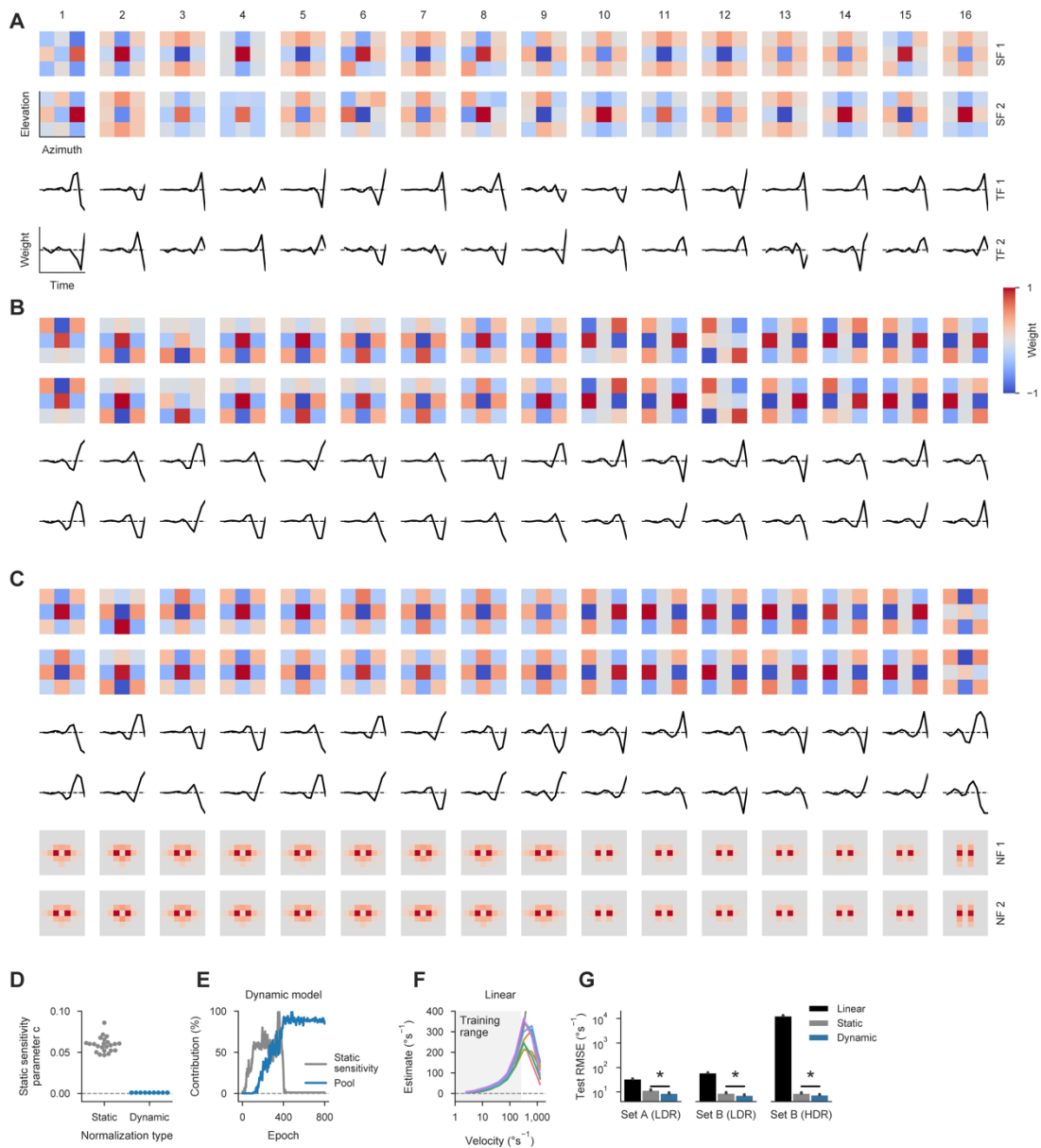


Figure S6 | Detailed receptive fields and performance data for task-driven model. Related to Figure 6.

(A–C) Receptive fields and temporal filters for 16 models of each non-linearity configuration (A, linear; B, static; C, dynamic). Models were sorted by test set error (increasing from left to right). Each pair of spatial and temporal filters was normalized to the maximum absolute weight across both channels (SF = spatial filter, TF = temporal filter, NF = normalization filter). Axis limits are the same as in Figure 6. (D) Values of sensitivity parameter c for all static (N=23) and dynamic (N=16) normalization models. (E) Evolution of weights for a single dynamic model. Both curves were independently normalized to their maximum across epochs. Pool contribution was quantified as the sum of weights across both 11 x 11 x 1 normalization filters. (F) Velocity tuning curves of best-performing linear model for various images (analogously to Figure 6G). Gray curve indicates true scene velocity on logarithmic axis. (G) Quantification of average model performance for all tested data sets (analogously to Figure 6H; LDR = low dynamic range, HDR = high dynamic range). See STAR Methods for details on how data sets were generated. Note that performance is plotted on a logarithmic axis. N=22/23/16 for linear/static/dynamic; *P<0.001; t=9.01/7.51/7.72 for set A/set B (LDR)/set B (HDR); Student’s t-test with assumed equal variance; only difference between static and dynamic was tested.

Cell type	L_{fg}	L_{bg}	p	c₅₀	w_{pool}	q	Norm. index	R²_{DivisiveNorm}	R²_{linear}
L1	1.47	0.07	1.10	0.53	0.22	0.97	0.42	98.39 ± 0.10	92.55 ± 0.14
L2	1.10	0.05	1.37	0.23	0.36	0.77	1.58	99.29 ± 0.03	85.17 ± 0.13
L3	1.68	0.16	1.46	1.00	0.00	1.27	0.00	95.90 ± 0.08	97.17 ± 0.07
L4	1.41	0.12	1.23	0.53	0.32	1.09	0.61	98.94 ± 0.04	93.71 ± 0.07
L5	1.04	0.05	1.29	0.14	0.19	1.10	1.36	94.51 ± 0.23	69.34 ± 0.24
Mi1	1.03	0.03	1.21	0.06	0.25	1.05	4.33	97.37 ± 0.14	56.26 ± 0.41
Mi4	1.61	0.33	0.90	1.00	0.31	5.92	0.31	90.08 ± 0.26	87.50 ± 0.35
Mi9	1.69	0.23	0.99	1.00	0.40	2.87	0.40	92.40 ± 0.24	89.61 ± 0.32
T4	0.96	0.01	2.47	0.11	0.49	0.74	4.45	96.78 ± 0.15	74.17 ± 0.35
T5	1.08	0.07	1.97	0.26	1.17	0.92	4.55	97.02 ± 0.13	77.27 ± 0.27
Tm1	0.98	0.09	1.87	0.18	0.86	0.71	4.75	97.53 ± 0.11	78.67 ± 0.29
Tm2	1.08	0.17	1.36	0.20	1.14	0.91	5.76	97.58 ± 0.08	73.09 ± 0.32
Tm3	1.02	0.01	1.97	0.16	0.53	0.72	3.39	97.97 ± 0.12	82.33 ± 0.20
Tm4	1.06	0.11	2.33	0.40	1.44	0.81	3.61	96.77 ± 0.16	76.96 ± 0.37
Tm9	1.83	0.50	0.92	0.98	1.01	1.65	1.03	96.37 ± 0.14	87.42 ± 0.25

Table S1 | Fits for divisive normalization model. Related to Figure 3.

Short name	Full genotype	Used in
WT	w ⁺ /w ⁺ ; +/+; +/+	Figure 1, Figure 2
T4/T5 block	w ⁺ /w ⁻ ; R59E08-AD/UAS-TNT; R42F06-DBD/+	Figure S2
L1-GCaMP6f	w ⁺ /w ⁻ ; VT027316-AD/UAS-GCaMP6f; R40F12-DBD/UAS-GCaMP6f	Figure 3, Figure S3–S4
L2-GCaMP6f	w ⁺ /w ⁻ ; R53G02-AD/UAS-GCaMP6f; R29G11-DBD/UAS-GCaMP6f	Figure 3, Figure S3–S4
L3-GCaMP6f	w ⁺ /w ⁻ ; R59A05-AD/UAS-GCaMP6f; R75H07-DBD/UAS-GCaMP6f	Figure 3, Figure S3–S4
L4-GCaMP6f	w ⁺ /w ⁻ ; R20A03-AD/UAS-GCaMP6f; R31C06-DBD/UAS-GCaMP6f	Figure 3, Figure S3–S4
L5-GCaMP6f	w ⁺ /w ⁻ ; R21A05-AD/UAS-GCaMP6f; R31H09-DBD/UAS-GCaMP6f	Figure 3, Figure S3–S4
Mi1-GCaMP6f	w ⁺ /w ⁻ ; R19F01-AD/UAS-GCaMP6f; R71D01-DBD/UAS-GCaMP6f	Figure 3, Figure 4, Figure S4
Tm3-GCaMP6f	w ⁺ /w ⁻ ; R13E12-AD/UAS-GCaMP6f; R59C10-DBD/UAS-GCaMP6f	Figure 3, Figure 4, Figure S4
Mi4-GCaMP6f	w ⁺ /w ⁻ ; R48A07-AD/UAS-GCaMP6f; R13F11-DBD/UAS-GCaMP6f	Figure 3, Figure S4
Mi9-GCaMP6f	w ⁺ /w ⁻ ; R48A07-AD/UAS-GCaMP6f; VT046779-DBD/UAS-GCaMP6f	Figure 3, Figure S4
Tm1-GCaMP6f	w ⁺ /w ⁻ ; R41G07-AD/UAS-GCaMP6f; R74G01-DBD/UAS-GCaMP6f	Figure 3, Figure 4, Figure S4
Tm2-GCaMP6f	w ⁺ /w ⁻ ; +/UAS-GCaMP6f; VT012282/UAS-GCaMP6f	Figure 3, Figure 4a-c, Figure S4
Tm2split-GCaMP6f	w ⁺ /w ⁻ ; R28D05-AD/UAS-GCaMP6f; R82F12-DBD/UAS-GCaMP6f	Figure 4
Tm4-GCaMP6f	w ⁺ /w ⁻ ; +/UAS-GCaMP6f; R35H01/UAS-GCaMP6f	Figure 3, Figure S4
Tm9-GCaMP6f	w ⁺ /w ⁻ ; +/UAS-GCaMP6f; VT065303/UAS-GCaMP6f	Figure 3, Figure S4
T4-GCaMP6f	w ⁺ /w ⁻ ; VT016255-AD/UAS-GCaMP6f; VT012314-DBD/UAS-GCaMP6f	Figure 3, Figure S3, Figure S4
T5-GCaMP6f	w ⁺ /w ⁻ ; VT013975-AD/UAS-GCaMP6f; R42H07-DBD/UAS-GCaMP6f	Figure 3, Figure S3, Figure S4
Mi1-GCaMP6f, TNT-E	w ⁺ /w ⁻ ; R19F01-AD/UAS-TNT-E; R71D01-DBD/UAS-GCaMP6f	Figure 5
Mi1-GCaMP6f, TNTin	w ⁺ /w ⁻ ; R19F01-AD/UAS-TNTin; R71D01-DBD/UAS-GCaMP6f	Figure 5

Tm3-GCaMP6f, TNT-E	w ⁺ /w ⁻ ; R13E12-AD/UAS-TNT-E; R59C10- DBD/UAS-GCaMP6f	Figure 5
Tm3-GCaMP6f, TNTin	w ⁺ /w ⁻ ; R13E12-AD/UAS-TNTin; R59C10- DBD/UAS-GCaMP6f	Figure 5
Tm1-GCaMP6f, TNT-E	w ⁺ /w ⁻ ; R41G07-AD/UAS-TNT-E; R74G01- DBD/UAS-GCaMP6f	Figure 5
Tm1-GCaMP6f, TNTin	w ⁺ /w ⁻ ; R41G07-AD/UAS-TNTin; R74G01- DBD/UAS-GCaMP6f	Figure 5
Tm2split-GCaMP6f, TNT-E	w ⁺ /w ⁻ ; R28D05-AD/UAS-TNT-E; R82F12- DBD/UAS-GCaMP6f	Figure 5
Tm2split-GCaMP6f, TNTin	w ⁺ /w ⁻ ; R28D05-AD/UAS-TNTin; R82F12- DBD/UAS-GCaMP6f	Figure 5

Table S2 | Genotypes and abbreviations. Related to Figures 1–5.

Supplemental Reference

- S1. Arenz, A., Drews, M.S., Richter, F.G., Ammer, G., and Borst, A. (2017). The temporal tuning of the *Drosophila* motion detectors is determined by the dynamics of their input elements. *Curr. Biol.* 27, 929–944.

1 Contrast Normalization Affects Response Time- 2 Course of Visual Interneurons

3 **Nadezhda Pirogova^{1,2*}, Alexander Borst¹**

4 ¹Department Circuits-Computation-Models, Max Planck Institute of Neurobiology, 82152 Martinsried,
5 Germany

6 ²Graduate School of Systemic Neurosciences, LMU Munich, 82152 Martinsried, Germany

7 *nadezhda.pirogova@bi.mpg.de

8 **Abstract**

9 In natural environments, light intensities and visual contrasts vary widely, yet neurons have a
10 limited response range for encoding them. Neurons accomplish that by flexibly adjusting their
11 dynamic range to the statistics of the environment via contrast normalization. The effect of
12 contrast normalization is usually measured as a reduction of neural signal amplitudes, but
13 whether it influences response dynamics is unknown. Here, we show that contrast
14 normalization in visual interneurons of *Drosophila melanogaster* not only suppresses the
15 amplitude but also alters the dynamics of responses when a dynamic surround is present. We
16 present a simple model that qualitatively reproduces the simultaneous effect of the visual
17 surround on the response amplitude and temporal dynamics by altering the cells' input
18 resistance and, thus, their membrane time constant. In conclusion, single-cell filtering
19 properties as derived from artificial stimulus protocols like white-noise stimulation cannot be
20 transferred one-to-one to predict responses under natural conditions.

21 **Introduction**

22 Due to the complexity of the natural environment, the sensory inputs animals receive can vary by several
23 orders of magnitude, ranging from the uniformity of an open field on an overcast evening to the stark

24 visual contrast of forest trees on a sunny day. The spectrum such inputs can encompass is broad but
25 individual neurons have only a limited response range to map these inputs onto, be it the graded
26 membrane potential or spike rates. Therefore, the range of the responses has to be used efficiently to
27 match the inputs [1]. For the visual system, that means adapting not only to the mean light intensity of
28 each of the scenes but also to the fluctuations of these intensities around the mean, i.e. the image
29 contrast. Neurons solve this problem by contrast adaptation: they dynamically adjust their contrast
30 sensitivity to the statistics of the current environmental conditions [2]. This way, the sensitivity of the
31 neuron increases at low contrast, making the neuron more responsive to small changes, and decreases
32 at high contrast so that large changes in the stimulus do not lead to response saturation [3 - 5]. Thus, the
33 amount of information about the stimulus contained in the response is maximized [1, 2, 6]. Contrast
34 adaptation in the vertebrate retina has also been shown to induce changes in the dynamics of the signal,
35 increasing the signal processing speed, making the signal more transient and, thus, improving the cell's
36 ability to encode fast temporal changes [3 - 5].

37 The fruit fly *Drosophila melanogaster* is able to successfully navigate a variety of environments with
38 variable visual statistics [7]. Motion vision is a critical sensory cue for the fruit flies' course control
39 system [8], where self-motion is estimated from the optic flow to respond to different environments
40 reliably and robustly [9]. *Drosophila's* visual system is well-studied, with a large genetic toolbox
41 allowing for interrogating individual neurons of interest. The visual system of *Drosophila* starts at the
42 retina and comprises four sequential neuropils, namely, lamina, medulla, lobula, and lobula plate (Fig
43 1A). Visual signals are processed retinotopically in the fruit fly's optic lobe. From lamina cells onwards,
44 processing runs in two parallel pathways: an ON pathway processes light increments, an OFF pathway
45 processes light decrements [10 - 12]. Neurons in the medulla have been previously described as either
46 transient or tonic, based on their filtering properties [13]. Medulla neurons themselves do not respond
47 selectively to the direction of visual motion but form the main inputs to the first direction-selective cells,
48 T4 in the ON pathway, and T5 in the OFF pathway [14 - 16]. Contrast normalization is present in the
49 early visual system of *Drosophila* and arises first in transient neurons of the medulla [17]: a high-
50 contrast grating presented in the visual surround outside the neuron's receptive field (RF) suppresses

51 the amplitude of the response to a local stimulus presented in the center of the cell's RF. Contrast
52 sensitivity is usually measured as the response of the cell as a function of the local contrast (Fig 1B).
53 When this measurement is taken with a grating moving in the surround, contrast normalization shifts
54 the response curve on a logarithmic contrast axis to the right. This shift adjusts the steep part of the
55 neuron's response curve, i.e. where the cell is the most sensitive, to the prevalent contrast in the
56 surround. Drews et al. (2020) [17] characterized the compressive, normalizing signal as fast, integrating
57 spatially over a large area and deriving from neural feedback. Functionally, it significantly improves
58 the robustness of motion detection in natural scenes [17].

59 **Fig 1. Contrast Normalization in Early Visual System of *Drosophila*.** (A) Schematic representation
60 of early stages of the motion detection circuit. Highlighted are contrast normalization-exhibiting
61 neurons that provide major input to T4 and T5 cells, the first direction-selective neurons in the ON and
62 OFF pathways, respectively. (B) Contrast normalization experimental protocol and contrast tuning
63 curves for different medulla neurons: Mi1 (n = 20/5), Tm3 (n = 21/8), Tm1 (n = 21/7), and Tm2 (n =
64 20/6). Shaded areas around the curves show bootstrapped 68% confidence intervals. Adapted from
65 Drews et al. (2020) [17].

66 To understand the mechanism by which T4 and T5 cells become selective for the direction of image
67 motion, it is important to know the filtering properties of each of their input cells. To characterize the
68 response dynamics of visual interneurons in *Drosophila*, the following artificial stimuli were commonly
69 employed in the past: a) gratings with defined spatial wavelength and contrast moving at various
70 velocities [14, 18, 19, 20]; b) moving edges of defined polarity [11, 21], i.e. either a dark edge on a
71 bright background or the other way around; c) white noise stimuli consisting of statistically independent
72 flickering pixels or bars [13, 22 - 24]; d) defined luminance pulses or steps placed in the center of the
73 RF of the cell [20, 25]. Importantly in the present context, these stimuli differ from each other with
74 respect to the amount of contrast present in the surround.

75 As was already shown, the contrast of the surround can strongly suppress the amplitude of the responses
76 to the center stimulus [17]. In this paper, we asked whether contrast normalization in the *Drosophila*
77 visual system has an influence not only on the amplitude of the response but also on its temporal

78 dynamics, as was shown for the vertebrate retina [3-5]. This way, contrast normalization would affect
79 the cell's filtering properties and, therefore, might influence its role in motion vision. We focused on
80 four transient medulla neurons that have previously shown to exhibit contrast normalization and to
81 provide major input signals to the first direction-selective cells in the ON and OFF pathways in
82 *Drosophila* [15, 16]. By combining artificial stimuli widely used to characterize neuronal responses in
83 *Drosophila melanogaster*, we aim to untangle the effects of the local and global stimulus profile on the
84 cells' filtering properties.

85 **Materials and Methods**

86 **Data and Code Availability**

87 Raw data from calcium imaging experiments, code to replicate the figures, and the modelling code are
88 available in the GitHub repository: https://github.com/nopirogova/paper_signal_time-course/

89 **Flies**

90 Flies were raised and kept on standard cornmeal-agar medium on a 12h light/12h dark cycle at 25°C
91 and 60% humidity. The genetically-encoded calcium indicator GCaMP6f [26] was expressed using the
92 Gal4-UAS system [27], resulting in the following genotypes:

93 L1>GCa6f: *w+*; *VT027316-AD/UAS-GCaMP6f*; *R40F12-DBD/UAS-GCaMP6f*

94 Mi1>GCa6f: *w+*; *R19F01-AD/UAS-GCaMP6f*; *R71D01-DBD/UAS-GCaMP6f*

95 Tm1>GCa6f: *w+*; *R41G07-AD/UAS-GCaMP6f*; *R74G01-DBD/UAS-GCaMP6f*

96 Tm2>GCa6f: *w+*; *R28D05-AD/UAS-GCaMP6f*; *R82F12-DBD/UAS-GCaMP6f*

97 Tm3>GCa6f: *w+*; *R13E12-AD/UAS-GCaMP6f*; *R59C10-DBD/UAS-GCaMP6f*

98 The transgenic fly line driving split-Gal4 expression in L1 cells is courtesy of A. Nern, Janelia Research
99 Campus; lines for Mi1 and Tm3 cells were generated and described in Strother et al. (2017) [20], for
100 Tm1 and Tm2 in Davis et al. (2020) [28].

101 **Two-photon Imaging**

102 **Fly Preparation**

103 The flies were taken 2-5 days after eclosion and prepared as previously described [14, 20]. In short,
104 flies were anesthetized on ice, their backs, legs, and wings were fixed onto an acrylic glass holder, and
105 the back of the head was exposed in a chamber filled with Ringer's solution. The cuticle behind the
106 right eye was cut with a fine hypodermic needle and removed, along with muscles and air sacks,
107 uncovering the optic lobe.

108 **Image Acquisition**

109 Two-photon imaging [29] was performed on a custom-built microscope as described by Maisak et al.
110 (2013) [14], controlled with the ScanImage software (version 5.1) in MATLAB [30]. Imaging was
111 performed at an acquisition rate of 11.8 Hz with an image resolution of 128×128 pixels. As described
112 by Arenz et al. (2013) [13], imaging stacks were automatically registered in a custom-written software
113 to correct for the movement of the brain.

114 Regions of interest (ROIs) were drawn manually on the average raw image to extract responses of
115 individual neurons. For ON pathway medulla cells, Mi1 and Tm3, ROIs were drawn in the medulla
116 layer M10, for the OFF pathway medulla neurons, Tm1 and Tm2, in lobula layer Lo1, and for the lamina
117 neuron, L1, in the medulla layer M1. Fluorescence changes ($\Delta F/F$ values) were then calculated using a
118 standard baseline algorithm over the ROI [31].

119 **Visual stimulation**

120 **Arena**

121 Stimuli were projected with 2 commercial micro-projectors (TI DLP Lightcrafter 3000) onto a custom-
122 built cylindrical arena, as previously described by Arenz et al. (2017) [13]. Stimuli covered 180° in
123 azimuth and 105° in elevation of the visual field of the fly.

124 The projectors had a refresh rate of 180 Hz (at 8-bit color depth), their medium brightness was set to
125 the value of 50 on an 8-bit grayscale, corresponding to a medium luminance of 55 ± 11 cd/m². Stimuli
126 were rendered using custom-written software in Python 2.7 and Panda3D framework.

127 **Gaussian Noise Stimulus**

128 The stochastic noise stimulus used to determine the location of a cell's receptive field was pre-rendered
129 and generated as previously described in Arenz et al. (2017) [13]. Briefly, the 3-minute-long stimulus
130 consisted of 64×52 pixels that covered the whole screen, one pixel corresponding to a visual angle of
131 around 2.8° . The intensities of each pixel were drawn from a Gaussian distribution at 100% contrast
132 and low-pass filtered using a Gaussian window with a standard deviation of approximately 90 ms,
133 corresponding to a binary noise with the temporal cut-off frequency of 1 Hz.

134 The response of an individual cell, as imaged within an ROI, was used to reconstruct the cell's
135 spatiotemporal receptive field as described by Arenz et al. (2017) [13]. The obtained coordinates of its
136 position on the screen were used to center the step stimulus as described below.

137 **Visual Stimulation of the Medulla Neurons**

138 The coordinates of a cell's receptive field, obtained from its response to the Gaussian noise, were used
139 to position the step stimulus. Before presenting the stimulus, we verified that the RF center was
140 sufficiently distant from the border of the screen so that a significant part of the surround of the stimulus
141 could be displayed.

142 The center of the stimulus comprised a 5° circular window, in which a 1-second-long step of luminance
143 was presented. The contrast polarity of the step corresponded to the preference of the cell [22], i.e. for
144 Mi1 and Tm3, the luminance of the center increased from 0% to 50%, 75%, or 100% during the step,
145 and vice versa, the luminance of the center decreased from 100% to 50%, 25%, or 0% for L1, Tm1, and
146 Tm2. The center window was surrounded by a 30° gray annulus (medium luminance), intended to cover
147 the surround of the receptive field of the cell and prevent the stimulus surround from leaking into the
148 cell's RF.

149 Beyond the annulus, one of the four surround conditions was shown, all with the same mean luminance.
150 Each surround condition was shown 3 s before the luminance step in the center and remained on screen
151 for 2 s after the step. In the cases when the surround condition was dynamic, the dynamics of the
152 surround started 1 s before the luminance step in the center. The four surround conditions were as
153 follows: 1. uniformly gray (contrast 0%); 2. 20°-wavelength stationary grating (contrast 100%); 2. 20°-
154 wavelength grating moving at 1 Hz (contrast 100%); 4. stochastic stimulus, (here, a binary noise with
155 the temporal cut-off frequency of 10 Hz, further properties as described in Gaussian Noise Stimulus).
156 All stimuli were repeated 3 times in randomized order to prevent adaptation to any stimulus features.

157 **Data Analysis**

158 Data analysis was performed offline using custom-written routines in Matlab and Python 2.7 and Python
159 3.7 (with the SciPy and OpenCV-Python Libraries).

160 **Data Evaluation**

161 Relative fluorescence change ($\Delta F/F$) was calculated using a standard baseline algorithm over an
162 individual ROI as described in Jia et al. (2010) [31]. Briefly, raw signal was smoothed with a Gaussian
163 window with FWHM of 1 s, the minima in a 90 second-long sliding window were extracted, and the
164 trace was smoothed with a Gaussian window with FWHM of 4 min, resulting in the dynamic baseline,
165 F_0 . To better compare the time-course of the signals to different stimuli, response curves were also
166 normalized to their maximum, i.e. divided by the peak value reached during the stimulation period.

167 To filter out the cells for which the variance in responses between trials was caused by movement
168 artifacts, a signal-to-noise ratio (SNR) criterion was introduced. Here, only the recordings, in which the
169 inter-trial variance was smaller than the average cell response, were taken, i.e. the standard deviation of
170 the averaged signal had to be at least 115% of the mean standard deviation over trials. On average, over
171 90% of all cells measured passed the SNR criterion.

172 **Data Visualization**

173 For all experiments, responses for each cell were averaged over trials, normalized to the cell's
174 maximum, and further averaged over the cells. Additionally, to illustrate the temporal dynamics of the
175 responses, responses of each cell type for every condition were normalized to the condition's maximum
176 for the respective cell type.

177 To visualize response traces when a cell responded identically to several stimuli, an artificial gap was
178 introduced by offsetting one of the responses vertically. In these cases, the figure legend specifies for
179 which conditions the offset was introduced.

180 **Statistical Tests**

181 Shaded areas around the response curves show bootstrapped 68% confidence intervals around the mean
182 (estimated as corresponding distribution percentiles after resampling the data 1,000 times). Sample sizes
183 are given in each figure legend. The experimenters were not blinded to genotypes or conditions during
184 data gathering and analysis.

185 **Modelling**

186 The model comprised three stages through which an input, a step function, was sequentially processed.

187 The input was a 1 s pulse at either full amplitude or an amplitude of 0.2.

188 The processing cascade was as follows: a (1) band-pass filter, followed by either a (2.1) static or a (2.2)
189 dynamic nonlinearity, followed by a (3) low-pass filter.

190 (1) The first stage of the model, a band-pass filter, simulated the response of a cell membrane to a
191 luminance step received in the center of its RF. The band-pass filter was constructed as a combination
192 of a low- and a high-pass filter (LP and HP) with $\tau_{LP} = 200$ ms and $\tau_{HP} = 300$ ms.

193 (2) At the second stage, the response was transferred through a divisive nonlinearity, as often used to
194 describe “contrast normalization” i.e. the saturating contrast dependency.

195 (2.1) The stationary nonlinearity was constructed as follows:

196 $y(t) = \frac{x(t)}{x(t) + k},$

197 where x was the input and y the output signal amplitude, and k the parameter controlling the amount
 198 of saturation. k values of 0.2 and 1.0 were used to simulate the static and dynamic surround
 199 experimental conditions, respectively.

200 (2.2) The dynamic nonlinearity was constructed as follows:

201 $y(t) = \frac{x(t-1)E_{exc} + y(t-1)C_{\Delta t}}{x(t-1) + k + C_{\Delta t}},$

202 where x was the input and y the output signal amplitude, E_{exc} was set to 1.0, $C_{\Delta t}$ to 100.0, and k was
 203 the parameter controlling the amount of saturation. k values of 0.2 and 1.0 were used to simulate the
 204 static and dynamic surround experimental conditions, respectively.

205 This corresponds to:

206 $V(t) = \frac{g_{exc}E_{exc} + V(t-1)\frac{C}{\Delta t}}{g_{exc} + g_{leak} + \frac{C}{\Delta t}},$

207 where g_{exc} is the conductance of the excitatory current, V is the membrane potential, E_{exc} is the reversal
 208 potential of the excitatory current, C is the membrane capacitance, and g_{leak} is the conductance of the
 209 leak current.

210 As derived from the equation for the electrically passive membrane, which states that the sum of all
 211 currents across the membrane equals zero:

212 $-C \frac{dV(t)}{dt} = g_{exc}(V(t) - E_{exc}) + g_{leak}(V(t) - E_{leak}).$

213 (3) At the last stage, to account for the calcium indicator (GCaMP6f) dynamics, the signal was
 214 processed through a low-pass filter with $\tau_{Ca} = 200$ ms.

215 Finally, the responses were normalized as described in “Data Evaluation” so that the model output was
216 comparable to the experimental data.

217 **Results**

218 We performed *in vivo* 2-photon calcium imaging from axon terminals of Mi1 and Tm3 neurons in layer
219 M10 of the medulla, and from axon terminals of Tm1 and Tm2 neurons in layer Lo1 of the lobula.
220 These recording sites correspond to the locations where the respective neurons synapse onto the
221 dendrites of the first direction-selective cells: Mi1 and Tm3 onto T4 cells, and Tm1 and Tm2 onto T5
222 cells (see Fig 1A).

223 As we were interested in the effects that contrast normalization has on single cells, our stimulus protocol
224 included a stochastic noise stimulus to determine the location of the cell’s RF. This location was then
225 used to project further stimuli to specific screen locations relative to the cell’s RF (Fig 2A).

226 **Fig 2. Dynamic Surround Affects Response Amplitude and Kinetics.** (A) Experimental procedure:
227 (1) white noise stimulus, (2) receptive field (RF) reconstruction from single-neuron calcium signals, (3)
228 experimental stimuli centered on RF. (B) Stimulus protocol. Luminance step in RF center with 4
229 surround conditions. (C) Average Tm3 response to luminance step in RF center with gray, stationary,
230 moving grating, and stochastic stimulus surround; $n = 23/6$. Luminance step happened during the gray-
231 shaded period. Left: Amplitudes of cell responses. Responses are normalized to the cell's maximum.
232 Shaded areas around the curves show bootstrapped 68% confidence intervals. Right: Kinetics of cell
233 responses. Responses during each condition are normalized to the condition's maximum. Artificial gap
234 is created between responses to static conditions for easier visualization. See also Fig S1.

235 All of our stimuli followed the same pattern: a 1-second luminance step with a 100% amplitude was
236 shown in a 5° circular window positioned in the center of the cell’s receptive field. This step consisted
237 of a luminance increase from 0 to 1 for Mi1 and Tm3, and a luminance decrease from 1 to 0 for Tm1
238 and Tm2. The window was surrounded by a 30° annulus at an intermediate luminance level that covered
239 most, and often the entirety, of the cell’s receptive field. The annulus ensured that the surround of the

240 stimuli was not spilling over into the receptive field of the cell avoiding a direct stimulation of the
241 neuron. The size of 30° of the annulus was based on the average full width at half maximum (FWHM)
242 for the imaged cells: this value was previously determined to be 29° for Mi1 cells, 12° for Tm3 cells,
243 27° for Tm1 cells, and 31° for Tm2 cells [13]. Covering the rest of the screen (hereafter referred to as
244 “surround”), sparing the center and the annulus, one of four different surround stimuli was presented
245 (Fig 2B): (i) uniformly gray, i.e. medium luminance, (ii) stationary grating at full contrast; (iii) 20° -
246 wavelength full-contrast grating moving at 1 Hz; (iv) full-contrast stochastic binary pixel noise. This
247 selection of stimulus conditions ensured that we had both a static (conditions i and ii) and a dynamic
248 (conditions iii and iv) surround. The two different dynamic surround conditions further differentiated
249 between temporal dynamics containing motion (moving grating in condition iii) and no motion
250 component (stochastic noise in condition iv).

251 **Effect of Visual Surround on Response Amplitude and Kinetics**

252 First, we wanted to probe the cells of interest to determine which features of the stimulus surround
253 affected the response of the cell to stimulation of its receptive field center. To this end, our stimulus
254 consisted of a luminance step in the center and one of the four conditions in the surround. Fig 2C shows
255 responses of Tm3 neurons to the four stimulus conditions. For all the surround conditions, the cells
256 responded to the luminance step with a fast signal increase that began decaying while the step was still
257 present. However, the amplitudes of the response and the kinetics of the signal decay varied depending
258 on the dynamics of the surround. Stationary grating in the surround had no visible effect on the signal,
259 making the responses to the two static surround conditions, uniformly gray and stationary grating,
260 virtually identical (Fig 2C). In the presence of a dynamic surround, however, the amplitude of the
261 response was dramatically reduced. For a stochastic noise surround, the maximum signal amplitude
262 only reached about 46% of the peak response value observed for static surround (for the values, see
263 Table S2). The signal was suppressed even more strongly by a moving grating in the surround, reaching
264 only about 20% of the peak response value observed for static surround.

265 In addition to suppressing the amplitude, the dynamic surround affected the kinetics of the responses
266 (Fig 2C, right panel). As best seen after normalizing cell responses to their peaks, Tm3 responses to the
267 two static conditions had virtually identical kinetics, with a rather slow decay of the signal during the
268 luminance step. In the presence of a dynamic surround, however, responses decayed visibly faster and,
269 by the end of the luminance step (arrows in Fig 2C), decreased to approximately 35% and 43% of their
270 peak amplitudes in the stochastic noise and moving grating surround conditions, respectively, compared
271 to only a decrease to 82% in the presence of static surround (for all values, see Table S1). Interestingly,
272 the temporal profiles of the responses in the two conditions with dynamic surround were virtually
273 identical, as were the signals in the two static surround conditions.

274 These phenomena, i.e. response amplitude suppression and faster response decay caused by the dynamic
275 but not the static surround, as well as extremely similar temporal profiles of the responses within the
276 two static and the two dynamic surround conditions, were also observed in Mi1 (Fig S1). Curiously, in
277 the specific case of Mi1, a certain level of response was present already before the luminance step and
278 was subsequently suppressed by the dynamic surround. We conclude that dynamic, but not static,
279 surround has an effect on both the response amplitude and its kinetics.

280 Because of the high level of similarity of response time courses within the static and dynamic surround
281 conditions, and as we were interested specifically in the effect of contrast normalization on the kinetics
282 of the signal, we focused on one stimulus per condition category. Thus, in further experiments, we used
283 the uniformly gray surround to represent the stationary and the moving grating for the dynamic surround
284 condition.

285 **Response Kinetics of Transient Medulla and Transmedulla Neurons**

286 With this stimulus protocol, we probed the other cells, i.e. Mi1, Tm1, and Tm2, that had previously
287 shown to exhibit contrast-normalizing properties [17]. Here, a similar picture arose (Fig 3). Firstly, all
288 the cells tested responded to the luminance step in the center of their receptive field, regardless of the
289 surround. Contrast normalization had a dramatic effect on the amplitude of the cells' responses, strongly
290 decreasing the size of the response in the presence of moving grating in the surround (Fig 3A). Here,

291 the response of Tm3 was suppressed the most: under dynamic surround conditions, its peak signal
292 corresponded to only about 20% of its response maximum in the static surround condition. For the rest
293 of the cells, the peak amplitude in the dynamic surround condition constituted approximately 58%
294 (Mi1), 58% (Tm1), and 45% (Tm2) of their highest signal observed in the uniformly gray surround
295 condition.

296 **Fig 3. Contrast Normalization Affects Response Amplitude and Kinetics.** Average responses of
297 contrast normalization-exhibiting neurons to luminance step in the RF center with gray and moving
298 grating surround. Mi1 (n = 98/25); Tm1 (n = 24/9); Tm2 (n = 22/7), and Tm3 (n = 65/16). Luminance
299 step happened during the gray-shaded period. (A) Amplitudes of cell responses. Responses are
300 normalized to the cell's maximum. Shaded areas around the curves show bootstrapped 68% confidence
301 intervals. (B) Kinetics of cell responses. Responses during each condition are normalized to the
302 condition's maximum.

303 The dynamic surround also had a pronounced effect on the temporal profile of the responses (Fig 3B),
304 with the signal decaying faster when a moving grating was present in the surround. Numerically, when
305 comparing the level of signal decay reached by the end of the luminance step in the two conditions, Mi1
306 response decreased to about 86% of its own maximum response amplitude in the dynamic condition,
307 compared to 97% in the static surround condition. For Tm3, the signal decayed to 43% of its own
308 maximum in the presence of a moving grating and to 82% in the presence of a uniformly gray surround.
309 For Tm1, these values for the dynamic and the static surround constituted 67% and 88%; and for Tm2,
310 22% and 35%, respectively.

311 We conclude that, in all the transient medulla cells that had previously been reported to exhibit contrast
312 normalization, the dynamic surround not only suppresses their response amplitude but also has a
313 pronounced effect on their temporal profile.

314 **Modelling Effects of Dynamic Surround**

315 How can we explain that a moving stimulus in the surround, outside the receptive field of a neuron,
316 affects both the amplitude and the time-course of the response? Classically, the phenomenon of contrast
317 normalization is attributed to a divisive nonlinearity. The saturating contrast dependency of the response
318 to a stimulus within the receptive field is well described by the following formula [2]:

$$319 \quad (1) \quad y(t) = \frac{x(t)}{x(t) + k}$$

320 with x being the input and y the output signal amplitude, and k the parameter which controls the amount
321 of saturation: for small values of k with respect to x , y strongly saturates with increasing x (Fig 4A).
322 This corresponds to the situation of a static surround. Conversely, for k being large with respect to x , y
323 grows in a rather linear way with increasing x . This corresponds to the situation where a grating is
324 moving in the surround.

325 **Fig 4. Model with Dynamic Nonlinearity Reproduces Dynamic Surround Effect on Response**
326 **Amplitude and Kinetics.** (A) Model schematic: input step is sequentially passed through band-pass
327 filter (BP), nonlinearity (NL), and low-pass filter (LP). NL with $k = 1$ corresponds to moving grating
328 surround condition, $k = 0.2$ to gray surround condition. (B) Model cascade with stationary nonlinearity.
329 (C) Model cascade with stationary nonlinearity and varying input amplitudes. (D) Model cascade with
330 dynamic nonlinearity and varying input amplitudes. (B-D) For easier visualization, an artificial gap is
331 created between model responses to moving grating and gray surround conditions.

332 Consider the following signal processing cascade (Fig 4A). It consists of a band-pass filter (Fig 4A,
333 left), the above-mentioned nonlinearity (Fig 4A, middle), and a final low pass filter (Fig 4A, right) to
334 account for the calcium indicator dynamics. The nonlinearity is shown for two values of k , 0.2 (blue,
335 representing static surround) and 1.0 (red, representing moving surround). If we stimulate the cell with
336 a 1s pulse of light (Fig 4B, left), the signal amplitude will be different after the nonlinearity, depending
337 on whether the surround is static or moving (Fig 4B, middle). However, the shape of the signal is also
338 slightly different, as can be seen after the output signals are normalized to their maximum amplitude

339 (Fig 4B, right). Thus, a stationary nonlinearity of the kind shown in formula (1) will affect not only the
340 signal amplitude but also the signal dynamic.

341 If this explains the phenomenon shown in Fig 3, we should obtain the same signal amplitude and time-
342 course for a static surround that we observed for a moving surround by simply reducing the stimulus
343 amplitude accordingly (Fig 4C). A stimulus amplitude of 0.2 for a static surround resulted in a signal
344 that is identical in amplitude and shape to the one resulting from a stimulus amplitude of 1.0 for a
345 moving surround (Fig 4C, middle to right).

346 There is, however, another explanation for the phenomenon taking into account the biophysics of a
347 neuron. The following equation describing an electrically passive membrane states that the sum of all
348 currents across the membrane equals zero:

$$349 \quad (2) \quad -C \frac{dV(t)}{dt} = g_{exc}(V(t) - E_{exc}) + g_{leak}(V(t) - E_{leak})$$

350 Here, $V(t)$ is the membrane potential, C the membrane capacitance, g_{exc} and E_{exc} the conductance
351 and the reversal potential of the excitatory current, and g_{leak} and E_{leak} the conductance and the reversal
352 potential of the leak current. Under steady-state conditions, i.e. $\frac{dV(t)}{dt} = 0$, and considering $V(t)$ relative
353 to E_{leak} (i.e. $E_{leak} = 0$), this equation becomes:

$$354 \quad (3) \quad V(t) = \frac{g_{exc}}{g_{exc} + g_{leak}} E_{exc}$$

355 With g_{exc} being the signal driving $V(t)$, the correspondence to equation (1), i.e. contrast normalization,
356 is obvious: g_{leak} becomes the factor controlling the amount of saturation. Under non-steady-state
357 conditions, however, g_{leak} , together with C , affects the membrane time-constant and thus the dynamic
358 of the membrane voltage. In other words: taking into account the biophysics of a neuron's membrane,
359 contrast normalization will alter the dynamic of the output signal, in addition to the effect described
360 above, by altering the membrane time constant.

361 Rewriting equation (2) as a difference equation results in the following equation:

362 (4)
$$V(t) = \frac{g_{exc}E_{exc} + V(t-1)\frac{C}{\Delta t}}{g_{exc} + g_{leak} + \frac{C}{\Delta t}}$$

363 With the capacitive current being small relative to the leak and excitatory current, equation (4)
 364 degenerates to equation (3), i.e. it describes the membrane voltage under steady-state conditions.

365 Can we still observe the same time-course for a small stimulus amplitude with a static surround and a
 366 large stimulus amplitude with a moving surround? As is shown in Fig 4D, the answer is ‘No’. Given a
 367 sufficiently large value of membrane capacitance, the two conditions cannot be interchanged with any
 368 combination of stimulus amplitudes: for both stimulus amplitudes, the responses with a moving
 369 surround are always faster than the responses with a static surround.

370 **Dynamic Nonlinearity in the Model Recapitulates Temporal Effects of Contrast**
 371 **Normalization**

372 In order to test which of the two explanations describes our data best, i.e. whether contrast normalization
 373 also affects the membrane time constant, we repeated the experiments on Tm3 and Mi1 cells, this time
 374 using three different stimulus amplitudes: 50%, 75%, and 100% (Fig 5A).

375 **Fig 5. Luminance Step Amplitude Has No Effect on Response Kinetics.** (A) Spatial and temporal
 376 stimulus profile. Left: 3 luminance step amplitudes with gray surround. Right: 3 luminance step
 377 amplitudes with moving grating surround. (B) Tm3 responses to 50%, 75%, and 100% luminance steps
 378 with gray surround; and to 50%, 75%, and 100% luminance steps with moving grating surround; n =
 379 21/4. Luminance steps happened during the gray-shaded period. Left: Amplitudes of cell responses.
 380 Responses are normalized to the cell's maximum. Shaded areas around the curves show bootstrapped
 381 68% confidence intervals. Right: Kinetics of cell responses. Responses during each condition are
 382 normalized to the condition's maximum. An artificial gap is created between responses to 100% and
 383 75% steps for easier visualization. (C) Mi1 responses to 50%, 75%, and 100% luminance step with gray
 384 surround; and to 50%, 75%, and 100% luminance step with moving grating surround; n = 26/8.
 385 Luminance step happened during the gray-shaded period. Left and right as in (B). See also Fig S2.

386 With decreasing input amplitude, the amplitude of the response also decreased, scaling, however, in a
387 non-linear way (Fig 5B and 5C). In Tm3, inputs of 75% and 100% in the dynamic surround condition
388 yielded virtually identical response amplitudes, while the amplitude of the response to a 50% luminance
389 step was significantly lower (Fig 5B, left). Mi1 exhibited similar behavior in the static surround
390 condition (Fig 5C, left).

391 However, inspection of the temporal profile of the responses (Fig 5B, C, right) reveals that the kinetics
392 of the signals remained the same within the static and the dynamic surround conditions, independent of
393 the input amplitude, with responses decaying visibly faster in the presence of a dynamic surround. Here,
394 by the end of the luminance step, Mi1 responses to all three input amplitudes decayed to about 85% of
395 the respective maximum if a moving grating was shown in the surround. In contrast, the responses
396 decreased to only 98% of the respective maximum if the surround was uniformly gray. Similarly, for
397 Tm3, the signal also decayed equally in response to the three stimuli, reaching approximately 39.5 -
398 43.5% of its maximum in the dynamic surround condition and roughly 80% if the surround was static.

399 Additionally, we used the same stimulus protocol and imaged the response of the lamina neuron L1
400 (Fig S2), previously shown not to exhibit contrast normalizing properties [17]. Here, the amplitude of
401 the response showed a clear dependence on the amplitude of the input step while the temporal kinetics
402 of the signal remained the same, independent of the properties of the surround or the amplitude of the
403 input.

404 Taken together, our results demonstrate that the amplitude of the input stimulus does not have a
405 pronounced effect on the temporal properties of the response. The results also show that the differences
406 in the signal time course are purely linked to the presence or absence of a dynamic surround. Therefore,
407 the change in the response dynamics cannot be explained as resulting from a static saturation
408 nonlinearity. Conversely, the results are in line with changes in the membrane time constant that depend
409 on the properties of the surround, thus supporting a model with a dynamic nonlinearity (Fig 4D).

410 **Discussion**

411 Characterizing the temporal filter properties of visual interneurons, we have shown that contrast
412 normalization has a large influence not only on the amplitudes of the cells' responses but also on their
413 temporal dynamics. This effect was present in all medulla neurons that had previously been shown to
414 possess contrast normalization properties (Fig 3) and was exerted only by dynamic, but not by stationary
415 stimuli present in the visual surround of the cells. We considered two models to explain this effect, one
416 with a stationary, the other with a dynamic nonlinearity (Fig 4). By demonstrating that the temporal
417 profile of the response was only influenced by the stimulus surround (Fig 5) and was not a simple
418 consequence of response saturation, we could exclude the model with a stationary nonlinearity. The
419 results were in line with our hypothesis that contrast normalization affects the membrane time constant.
420 The results, thus, favor the model with a dynamic nonlinearity.

421 In Drews et al. (2020) [17], the presence of contrast normalization in *Drosophila's* optic lobe and its
422 suppressive effect on the gain of the response has been established. Here, we demonstrate the effect of
423 contrast normalization on the temporal kinetics of the responses. Our results parallel the findings from
424 the vertebrate retina, where cells' processing has also been shown to speed up, leading to their responses
425 becoming more transient at higher contrasts [3-5]. At the same time, our findings seem to be, at first
426 sight, hard to reconcile with a study by Matulis et al. (2020) [32] who investigated temporal contrast
427 adaptation in the *Drosophila* optic lobe. They found different cells to be affected by the visual contrast
428 than the current study as well as what was described in Drews et al. (2020) [17]. Furthermore, in
429 electrophysiological recordings, contrast adaptation had no effect on the temporal dynamics of the
430 signal, while, in two-photon calcium recordings from the same cells' dendrites, higher contrasts were
431 found to slow the signals down - a finding that is opposite to the results from the vertebrate retina [3-5]
432 and from our study. Whatever the explanation of the latter differences between voltage and calcium
433 signals may be, these discrepancies may be partly attributed to the differences in the stimuli applied. In
434 order to study contrast adaptation, i.e. the change in the cell's signal in response to the contrast of the
435 entire visual scene, Matulis et al. (2020) [32] used full-field stimuli consisting of stochastic binary noise
436 that switched between periods of low and high contrast. Hence, the change in stimulus contrast applied

437 to both the cells' receptive field as well as the overall surround simultaneously. Our goal was to study
438 contrast normalization, i.e. the influence of the surround stimulus on the response of the cell to the
439 stimulus within its receptive field. Therefore, we as well as Drews et al (2020) [17] designed the
440 stimulus such as to stimulate the cells' receptive field and the surround separately. Taken together, the
441 effects of contrast normalization on the amplitude and the dynamics of the response to stimuli delivered
442 within the receptive field of the neurons, as found in Drews et al. and in this paper, seem to be distinct
443 from the effects of contrast adaptation, as found by Matulis et al.(2020) [32].

444 **Implications for Characterizing Filter Properties of Visual Interneurons**

445 In order to characterize the filter properties of single cells, a variety of different artificial stimuli are
446 typically used, amongst them stochastic pixel noise, bars, and gratings, to name only a few. These
447 stimuli offer numerous advantages for systematic studies: they are easy to parametrize by their mean
448 luminance, contrast, and spatial wavelength, and the resulting response is readily evaluated in terms of
449 the cell's impulse response, receptive field size, and frequency spectrum. However, the visual system
450 has evolved as an adaption to complex natural images. Therefore, the results obtained from the
451 simplified artificial stimuli may not be transferable to a cell's responses to naturalistic stimuli. Unlike
452 natural environments, many of the stimuli used to characterize response properties of visual
453 interneurons are tailored to the receptive field of the cell under study. However, as we have shown
454 above, the presence or absence of dynamic stimuli in the visual surround, far outside of the receptive
455 field of a cell, has a pronounced influence not only on the response amplitude but also on the temporal
456 response properties of the cell.

457 This has immediate implications for the interpretation of results obtained from visual interneurons in
458 general, and for neurons involved in motion vision in particular. Here, the various models proposed to
459 account for this computation [33-34], although different in detail, all share the principle that a direction-
460 selective output is achieved by a non-linear operation performed on differentially filtered signals
461 derived from adjacent image points. In the fly visual system, T4 and T5 cells are known to be the first
462 direction-selective neurons [14]. While contrast normalization has indeed been demonstrated in a

463 number of columnar neurons providing input to T4 and T5 cells [17], their specific contributions to
464 motion vision were derived from their temporal filtering properties which, in turn, were determined
465 using stochastic pixel noise [13]. However, depending on the specific cell under study, the temporal
466 response properties might be quite different whether the cell is stimulated by a full-field grating or by
467 a local contrast step. To avoid this dilemma and in order to feed the model with faithful signals, input
468 neurons should ideally be measured under stimulus conditions identical to those used to characterize
469 the motion-sensitive neurons. Interestingly, this exact path has been chosen in a recent study on T4 cells
470 [35].

471 **Functional Consequences of Contrast Normalization**

472 We focused on the visual interneurons with band-pass filtering properties [13], that provide excitatory
473 input onto direction-selective T4 and T5 cells [16], exhibit contrast normalization [17], and are often
474 used as input signals of various motion detector model simulations [13, 23]. Incorporating spatial
475 contrast normalization into correlation-based models of motion vision has already been demonstrated
476 to drastically improve the models' performance [17]. Here, we show that, in addition to the suppressive
477 effect on the amplitude of the response, contrast normalization also alters the temporal dynamics of the
478 signal and, thus, the filtering properties of the neurons. This change in the time course of the inputs has
479 potential implications for the temporal dynamics of T4 and T5 neurons by adjusting the cells' speed
480 tuning to the prevailing speed in the surround. To determine the influence that visual surround has on
481 temporal filtering properties and velocity tuning of the neurons, the motion vision circuit needs to be
482 scanned with a set of dynamic global stimuli at varying frequencies.

483 Contrast normalization does not alter the amplitude and the temporal dynamics of all the cells it affects
484 by an equal amount. Consequently, the relative contribution of the different inputs to the T4 and T5
485 cells will vary depending on the global structure of the stimulus. In the ON pathway, in the presence of
486 a global visual surround, Tm3 responses are suppressed and sped up to a higher degree than the
487 responses of Mi1 neurons (Fig 3). As these two cells constitute a large portion of T4 input [16, 36], the
488 unequal effects of contrast normalization, depending on the presence of a local or a global stimulus,

489 would skew the ratio of the neurons' outputs onto T4. To emphasize, this has severe implications for
490 the resulting output of the computation because the composition of the inputs varies, depending on the
491 profile of the visual stimulus.

492 **Mechanism of Contrast Normalization**

493 The contrast normalization mechanism relies on neural feedback. At least part of this feedback comes
494 from one or more medulla neurons [17]. The feedback does not solely rely on the output of the neuron
495 itself, as blocking the cell's output is not enough to completely abolish contrast normalization of its
496 response.

497 Even though the normalizing cell has not yet been identified, the search can be focused, as we can
498 determine a number of characteristics that describe the feedback cell. Firstly, the normalizing feedback
499 neuron has temporal band-pass filtering properties, as a stimulus with a high-contrast but static surround
500 elicits the same response as a purely local stimulus (Fig 2). Secondly, the normalizing feedback cell is
501 not selective for the direction of motion, as a surround with purely temporal dynamics is enough to
502 reproduce the normalizing effects of a global stimulus that contains spatial motion. Thirdly, the
503 normalizing feedback neuron either has a large receptive field or is a part of an interconnected network
504 of the same cell type with smaller receptive fields, as the strength of the normalization increases as the
505 surround grows in size, extending well beyond 50° [17].

506 We also make a clear prediction about the mechanism, via which the normalization is achieved in the
507 neuron that exhibits the contrast normalization properties. The effects of contrast normalization on
508 amplitude and kinetics of the signal are not a result of a simple saturation (Fig 4), instead, we can
509 reproduce these effects by altering the cell's input resistance. We hypothesize that the input resistance
510 of the cell should drop significantly in the presence of a global stimulus with a dynamic surround.
511 However, measuring such a drop in the input resistance in the presence of the global dynamic surround
512 at varying speeds might be a complicated experiment to perform. The contrast normalization effect
513 described here is seen when imaging at the axon terminals of the cells, at the site of their synaptic output
514 onto the first direction-selective cells, i.e. T4 and T5 cells. In the ON pathway, this corresponds to

515 medulla layer M10. Input resistance measurements of these cells, however, are only possible at the
516 soma, which, in the case of ON-pathway interneurons, is located outside of the medulla. Supposing that
517 the neuronal computations are compartmentalized and local to the axons and dendrites, the drop in the
518 input resistance might not reach the soma and thus not be measurable there.

519 To uncover the physiological mechanism underlying contrast normalization, one could start by focusing
520 on octopamine, as the application of the octopamine receptor agonist chlordimeform (CDM) has been
521 shown to significantly speed up response of *Drosophila* visual interneurons [17], and stimulus-
522 dependence has been demonstrated to elicit changes in the shape of the response similar to those
523 produced by octopamine [24]. The effect we observed here might be of a similar nature, mimicking the
524 shift toward higher frequencies that occurs in the cells when the fly is in active locomotion [18], a state,
525 in which it receives an abundance of global visual cues from the environment. Exploring the connection
526 between the activation of octopamine receptors and the drop in the input resistance of the neuron in the
527 presence of a global dynamic stimulus might shed further light on the mechanism of contrast
528 normalization.

529 In summary, we built on the work in contrast normalization in the motion vision circuitry of *Drosophila*
530 *melanogaster* and demonstrated the dramatic effect that contrast normalization has on the temporal
531 characteristics of the interneurons' responses. Our findings illustrate the limitations of using simplified
532 artificial stimuli with varying spatial profiles to probe the filtering properties of single cells and the
533 constraints in transferring these results onto naturalistic conditions.

534 **Supporting Information**

535 **S1 Fig. Dynamic Surround Affects Response Amplitude and Kinetics.** Mi1 responses to a luminance
536 step in the RF center with gray, stationary, moving grating, and stochastic stimulus surround; n = 48/11.
537 Luminance step occurred during the gray-shaded period. Left: Amplitudes of cell responses. Responses
538 are normalized to the cell's maximum. Shaded areas around the curves show bootstrapped 68%
539 confidence intervals. Right: Kinetics of cell responses. Responses during each condition are normalized
540 to the condition's maximum.

541 **S2 Fig. Dynamic Surround and Input Amplitude Do Not Affect L1 Response Kinetics.** (A)
542 Schematic representation of early stages of the motion detection circuit. Highlighted is a lamina neuron
543 in the ON pathway that doesn't exhibit contrast normalization properties and provides major input to
544 contrast normalization-exhibiting neurons. (B) Contrast tuning curves for lamina neuron L1 (n = 14/4).
545 Shaded areas around the curves show bootstrapped 68% confidence intervals. Adapted from Drews et
546 al. (2020) [17]. (C) Spatial and temporal stimulus profile. Left: 3 luminance step amplitudes with gray
547 surround. Right: 3 luminance step amplitudes with moving grating surround. (D) L1 responses to
548 luminance steps of different amplitudes with gray and moving grating surround. Luminance step
549 happened during gray-shaded period. Left: Amplitudes of cell responses. Responses are normalized to
550 cell's maximum. Shaded areas around the curves show bootstrapped 68% confidence intervals. Right:
551 Response kinetics. Responses during each condition are normalized to condition's maximum.

552 **S1 Table.** Signal Amplitude after Luminance Step as a Percentage of Peak Amplitude during the Step.

	Mi1	Tm1	Tm2	Tm3
Stationary grating	98.58%	90.89%	32.52%	81.32%
Moving grating	86.09%	67.05%	22.19%	43.47%
Moving grating, 75% step	89.23%	N/A	N/A	40.23%
Moving grating, 50% step	84.57%	N/A	N/A	39.46%
Stochastic	83.96%	N/A	N/A	35.21%
Gray	96.91%	88.38%	35.24%	82.15%
Gray, 75% step	97.49%	N/A	N/A	81.89%
Gray, 50% step	98.52%	N/A	N/A	75.16%

553

554 **S2 Table.** Maximum Signal Amplitude as a Percentage of Peak Amplitude during Gray Condition.

	Mi1	Tm1	Tm2	Tm3
Stationary grating	99.89%	104.41%	101.06%	100.17%
Moving grating	58.49%	58.34%	45.19%	20.54%
Moving grating, 75% step	48.38%	N/A	N/A	17.46%

Moving grating, 50% step	33.29%	N/A	N/A	9.59%
Stochastic	65.43%	N/A	N/A	45.68%
Gray	100.00%	100.00%	100.00%	100.00%
Gray, 75% step	101.07%	N/A	N/A	84.26%
Gray, 50% step	87.85%	N/A	N/A	52.14%

555

556 **Acknowledgments**

557 We thank Jurgen Haag and Lukas Groschner for commenting on drafts of the manuscript. We thank
558 Wolfgang Essbauer and Michael Sauter for fly husbandry. This work was supported by the Deutsche
559 Forschungsgemeinschaft (SFB 870) and the Max Planck Society. N.P. and A.B. are members of the
560 Graduate School of Systemic Neurosciences (GSN) Munich.

561 **Author Contributions**

562 Conceptualization: Nadezhda Pirogova, Alexander Borst.

563 Data curation: Nadezhda Pirogova.

564 Formal analysis: Nadezhda Pirogova, Alexander Borst.

565 Funding acquisition: Alexander Borst.

566 Investigation: Nadezhda Pirogova.

567 Methodology: Nadezhda Pirogova, Alexander Borst.

568 Project administration: Nadezhda Pirogova.

569 Supervision: Alexander Borst.

570 Validation: Nadezhda Pirogova.

571 Visualization: Nadezhda Pirogova.

572 Writing – original draft: Nadezhda Pirogova.

573 Writing – review & editing: Alexander Borst.

574 **References**

575 [1] Laughlin S. A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch C*.
576 1981;36: 910–912.

577 [2] Carandini M, Heeger DJ. Normalization as a canonical neural computation. *Nat Rev Neurosci*.
578 2011;13: 51–62.

579 [3] Chander D, Chichilnisky EJ. Adaptation to temporal contrast in primate and salamander retina. *J*
580 *Neurosci*. 2001;21: 9904–9916.

581 [4] Kim KJ, Rieke F. Temporal contrast adaptation in the input and output signals of salamander
582 retinal ganglion cells. *J Neurosci*. 2001;21: 287–299.

583 [5] Baccus SA, Meister M. Fast and slow contrast adaptation in retinal circuitry. *Neuron*. 2002;36:
584 909–919.

585 [6] Smirnakis SM, Berry MJ, Warland DK, Bialek W, Meister M. Adaptation of retinal processing to
586 image contrast and spatial scale. *Nature*. 1997. pp. 69–73. doi:10.1038/386069a0

587 [7] Brinkworth RSA, O'Carroll DC. Robust models for optic flow coding in natural scenes inspired by
588 insect biology. *PLoS Comput Biol*. 2009;5: e1000555.

589 [8] Leonte M-B, Leonhardt A, Borst A, Mauss AS. Aerial course stabilization is impaired in motion-
590 blind flies. *J Exp Biol*. 2021;224. doi:10.1242/jeb.242219

591 [9] Borst A. Fly visual course control: behaviour, algorithms and circuits. *Nat Rev Neurosci*. 2014;15:
592 590–599.

593 [10] Joesch M, Schnell B, Raghu SV, Reiff DF, Borst A. ON and OFF pathways in *Drosophila* motion
594 vision. *Nature*. 2010;468: 300–304.

595 [11] Clark DA, Bursztyn L, Horowitz MA, Schnitzer MJ, Clandinin TR. Defining the computational
596 structure of the motion detector in *Drosophila*. *Neuron*. 2011;70: 1165–1177.

597 [12] Strother JA, Nern A, Reiser MB. Direct observation of ON and OFF pathways in the *Drosophila*
598 Visual System. *Current Biology*. 2014. pp. 976–983. doi:10.1016/j.cub.2014.03.017

599 [13] Arenz A, Drews MS, Richter FG, Ammer G, Borst A. The temporal tuning of the *Drosophila*
600 Motion Detectors Is Determined by the Dynamics of Their Input Elements. *Curr Biol.* 2017;27: 929–
601 944.

602 [14] Maisak MS, Haag J, Ammer G, Serbe E, Meier M, Leonhardt A, et al. A directional tuning map of
603 *Drosophila* elementary motion detectors. *Nature.* 2013;500: 212–216.

604 [15] Takemura S-Y, Nern A, Chklovskii DB, Scheffer LK, Rubin GM, Meinertzhagen IA. The
605 comprehensive connectome of a neural substrate for “ON” motion detection in *Drosophila*. *eLife.*
606 2017. doi:10.7554/elife.24394

607 [16] Shinomiya K, Huang G, Lu Z, Parag T, Xu CS, Aniceto R, et al. Comparisons between the ON-
608 and OFF-edge motion pathways in the *Drosophila* brain. *Elife.* 2019;8. doi:10.7554/eLife.40025

609 [17] Drews MS, Leonhardt A, Pirogova N, Richter FG, Schuetzenberger A, Braun L, et al. Dynamic
610 signal compression for robust motion vision in flies. *Curr Biol.* 2020;30: 209–221.e8.

611 [18] Suver MP, Mamiya A, Dickinson MH. Octopamine neurons mediate flight-induced modulation of
612 visual processing in *Drosophila*. *Current Biology.* 2012. pp. 2294–2302.
613 doi:10.1016/j.cub.2012.10.034

614 [19] Tuthill JC, Nern A, Holtz SL, Rubin GM, Reiser MB. Contributions of the 12 neuron classes in the
615 fly lamina to motion vision. *Neuron.* 2013. pp. 128–140. doi:10.1016/j.neuron.2013.05.024

616 [20] Strother JA, Wu S-T, Wong AM, Nern A, Rogers EM, Le JQ, et al. The emergence of directional
617 selectivity in the visual motion pathway of *Drosophila*. *Neuron.* 2017;94: 168–182.e10.

618 [21] Serbe E, Meier M, Leonhardt A, Borst A. Comprehensive characterization of the major
619 presynaptic elements to the *Drosophila* OFF motion detector. *Neuron.* 2016;89: 829–841.

620 [22] Behnia R, Clark DA, Carter AG, Clandinin TR, Desplan C. Processing properties of ON and OFF
621 pathways for *Drosophila* motion detection. *Nature.* 2014;512: 427–430.

622 [23] Gonzalez-Suarez AD, Zavatone-Veth JA, Chen J, Matulis CA, Badwan BA, Clark DA.
623 Manipulating neural dynamics to tune motion detection. *bioRxiv.* 2021. p. 2021.11.02.466844.
624 doi:10.1101/2021.11.02.466844

625 [24] Kohn JR, Portes JP, Christenson MP, Abbott LF, Behnia R. Flexible filtering by neural inputs
626 supports motion computation across states and stimuli. *Curr Biol.* 2021;31: 5249–5260.e5.

627 [25] Haag J, Mishra A, Borst A. A common directional tuning mechanism of *Drosophila* motion-
628 sensing neurons in the ON and in the OFF pathway. *Elife.* 2017;6. doi:10.7554/eLife.29044

629 [26] Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, et al. Ultrasensitive
630 fluorescent proteins for imaging neuronal activity. *Nature*. 2013;499: 295–300.

631 [27] Brand AH, Perrimon N. Targeted gene expression as a means of altering cell fates and
632 generating dominant phenotypes. *Development*. 1993;118: 401–415.

633 [28] Davis FP, Nern A, Picard S, Reiser MB, Rubin GM, Eddy SR, et al. A genetic, genomic, and
634 computational resource for exploring neural circuit function. *Elife*. 2020;9. doi:10.7554/eLife.50901

635 [29] Denk W, Strickler JH, Webb WW. Two-photon laser scanning fluorescence microscopy. *Science*.
636 1990;248: 73–76.

637 [30] Pologruto TA, Sabatini BL, Svoboda K. ScanImage: flexible software for operating laser scanning
638 microscopes. *Biomed Eng Online*. 2003;2: 13.

639 [31] Jia H, Rochefort NL, Chen X, Konnerth A. In vivo two-photon imaging of sensory-evoked dendritic
640 calcium signals in cortical neurons. *Nat Protoc*. 2011;6: 28–35.

641 [32] Matulis CA, Chen J, Gonzalez-Suarez AD, Behnia R, Clark DA. Heterogeneous temporal contrast
642 adaptation in *Drosophila* direction-selective circuits. *Curr Biol*. 2020;30: 222–236.e6.

643 [33] Hassenstein B, Reichardt W. Systemtheoretische Analyse der Zeit-, Reihenfolgen- und
644 Vorzeichenauswertung bei der Bewegungspertzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift für*
645 *Naturforschung B*. 1956. pp. 513–524. doi:10.1515/znb-1956-9-1004

646 [34] Barlow HB, Levick WR. The mechanism of directionally selective units in rabbit's retina. *J Physiol*.
647 1965;178: 477–504.

648 [35] Groschner LN, Malis JG, Zuidinga B, Borst A. A biophysical account of multiplication by a single
649 neuron. *Nature*. 2022;603: 119–123.

650 [36] Ammer G, Leonhardt A, Bahl A, Dickson BJ, Borst A. Functional specialization of neural input
651 elements to the *Drosophila* ON motion detector. *Curr Biol*. 2015;25: 2247–2253.

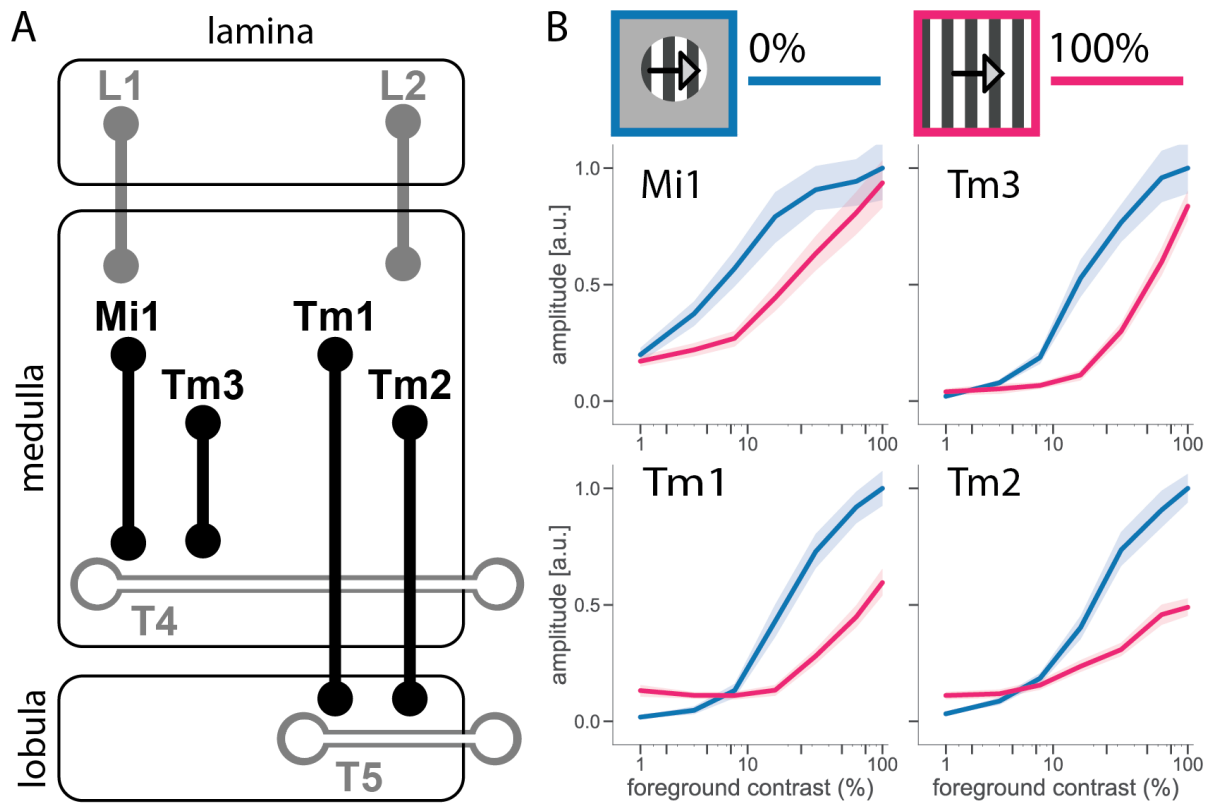


Fig 1. Contrast Normalization in Early Visual System of *Drosophila*.

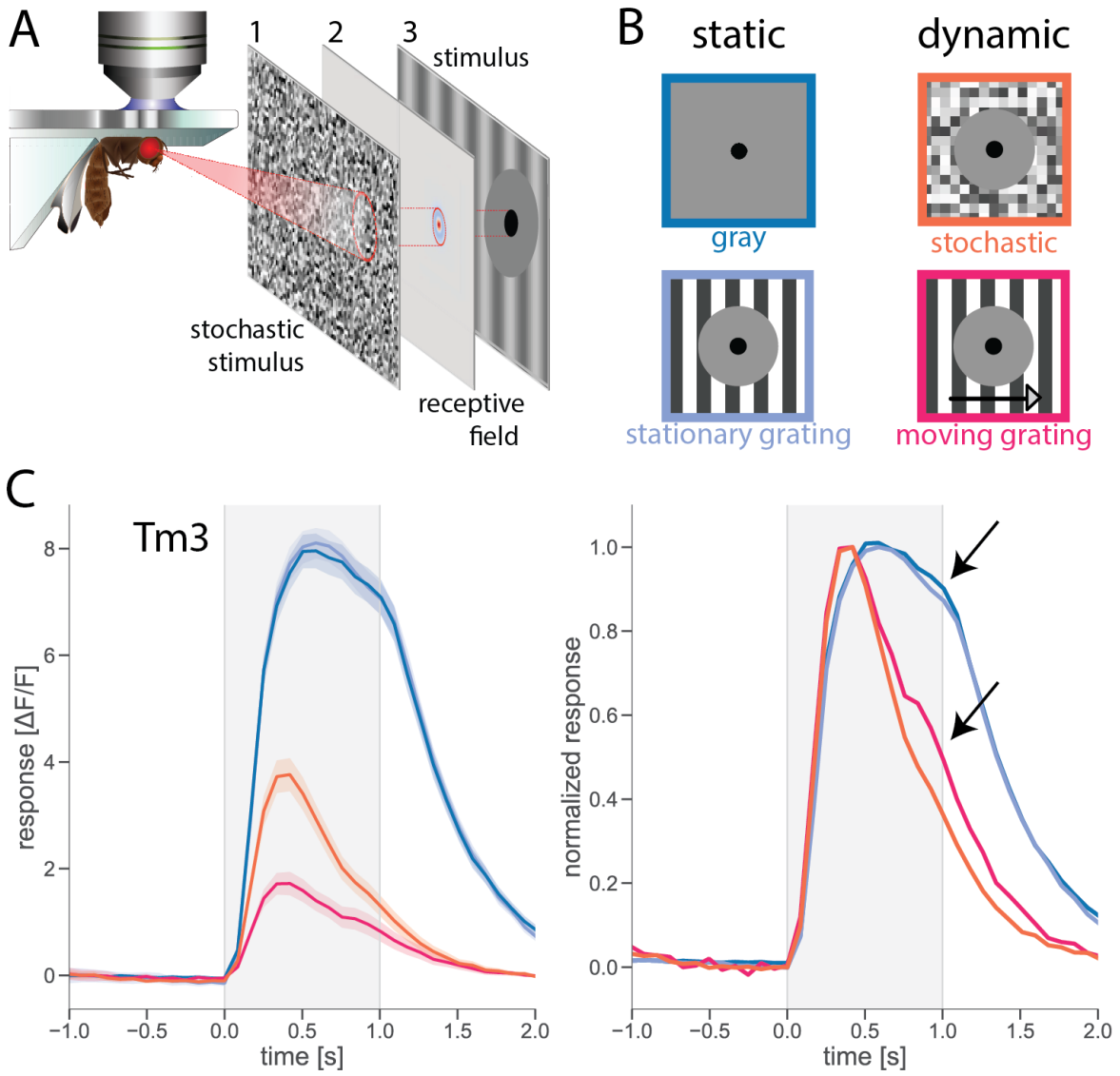


Fig 2. Dynamic Surround Affects Response Amplitude and Kinetics.

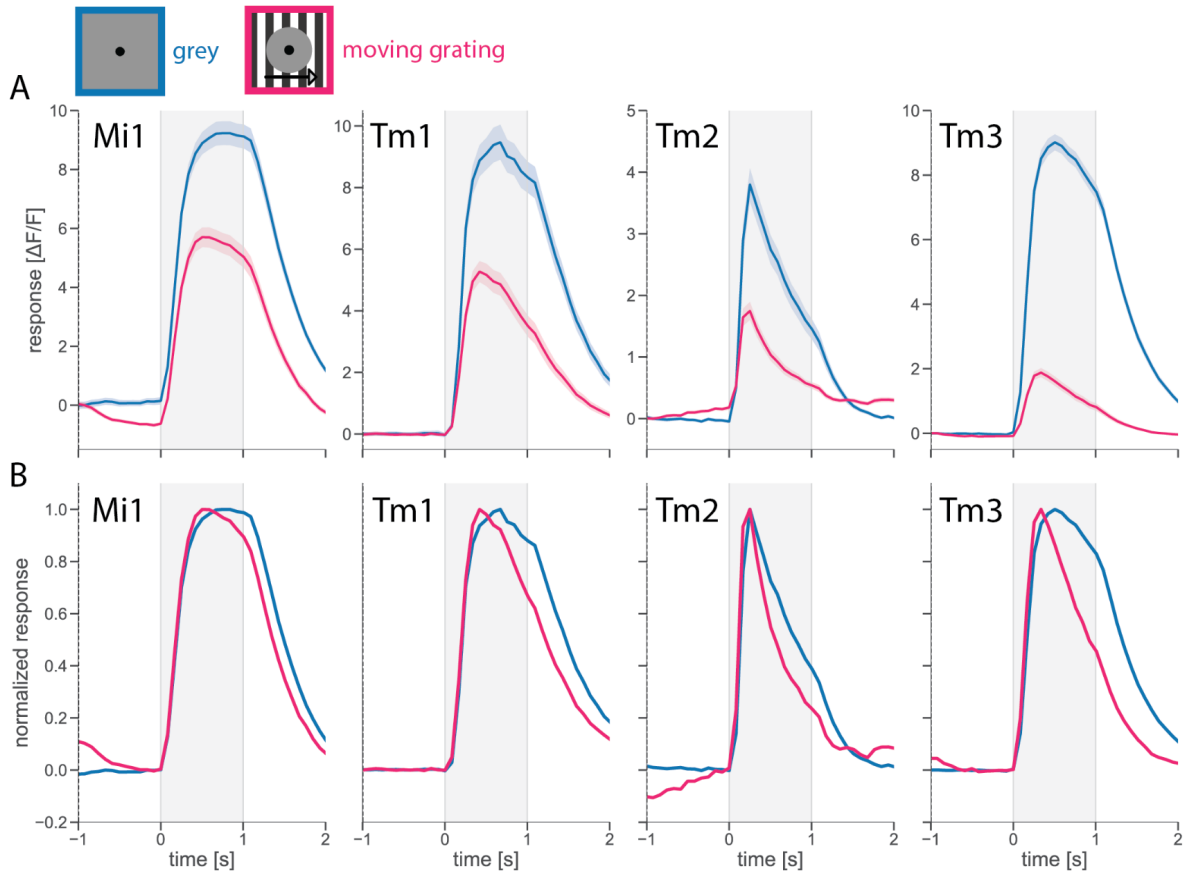


Fig 3. Contrast Normalization Affects Response Amplitude and Kinetics.

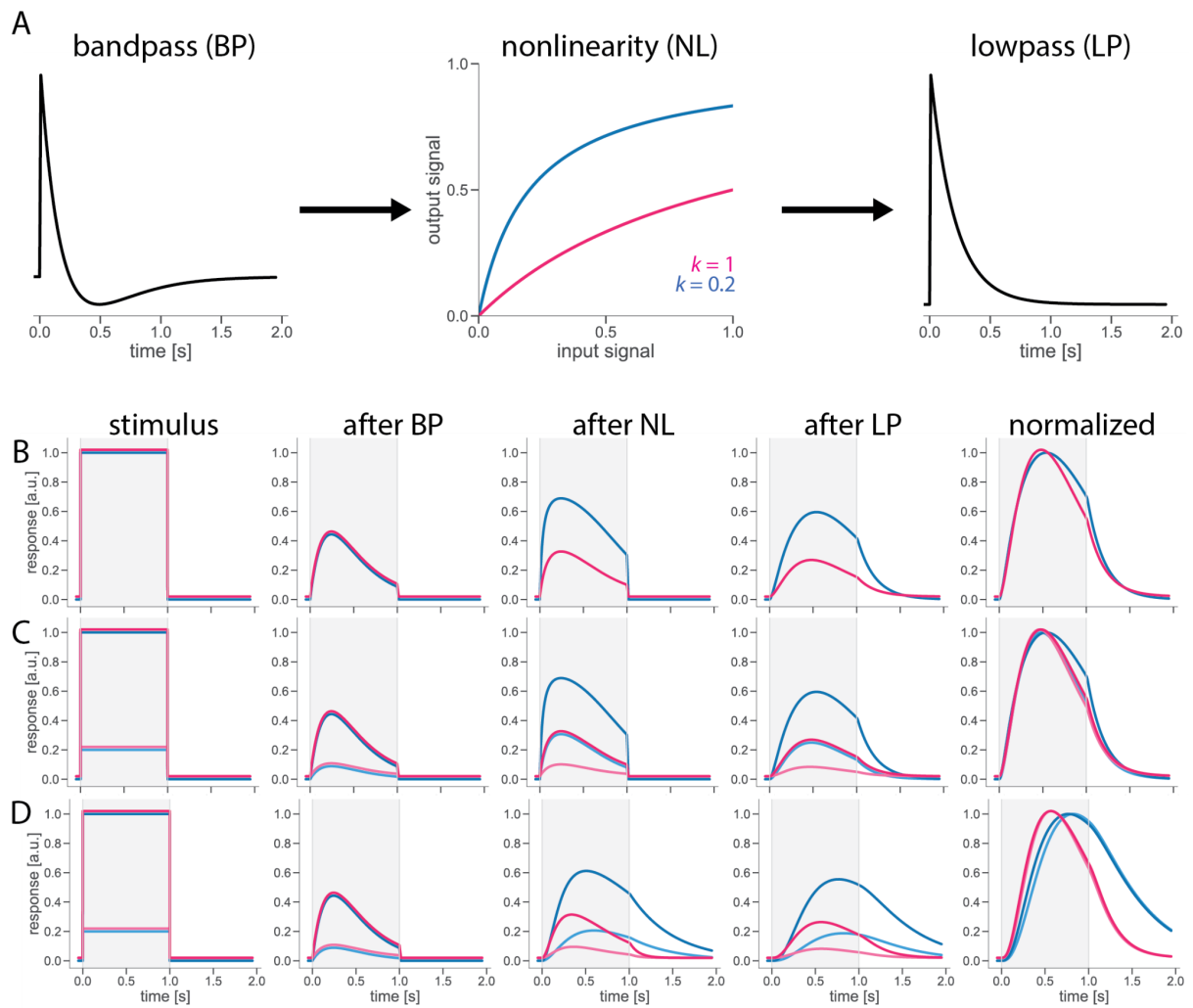


Fig 4. Model with Dynamic Nonlinearity Reproduces Dynamic Surround Effect on Response Amplitude and Kinetics.

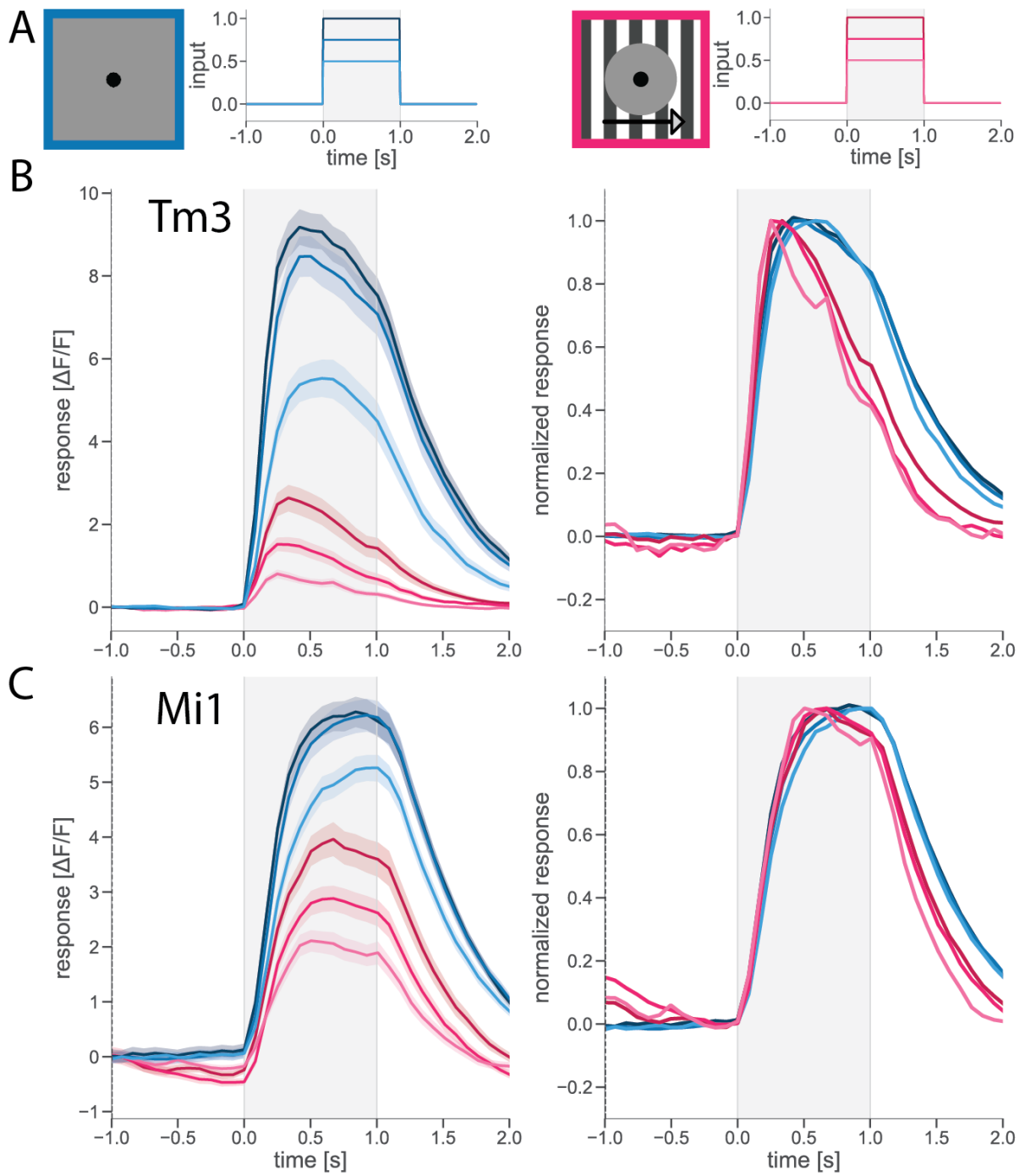
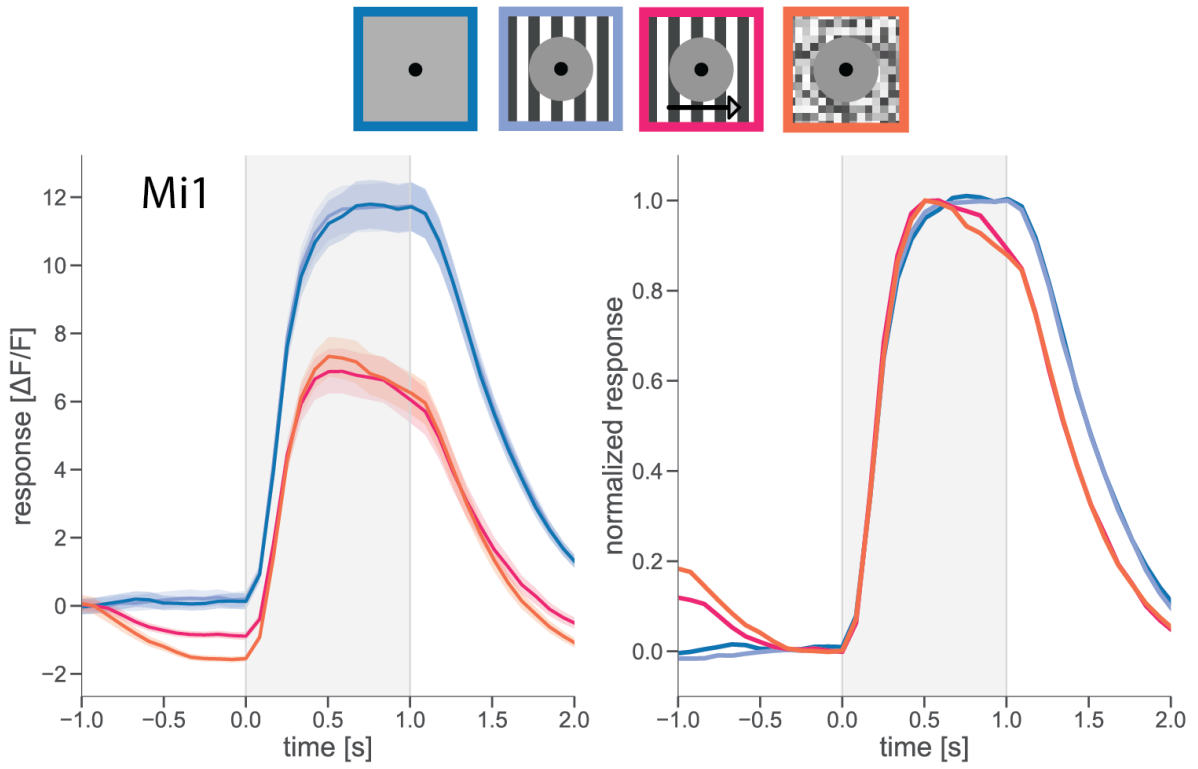
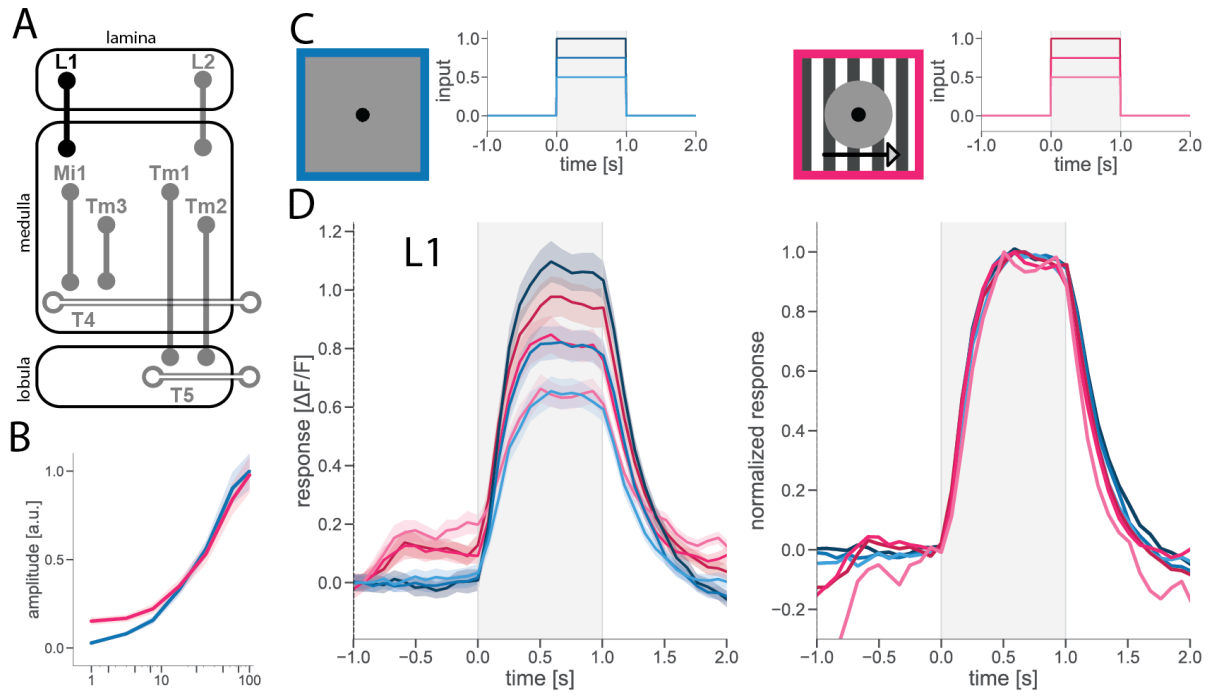


Fig 5. Luminance Step Amplitude Has No Effect on Response Kinetics.



S1 Fig. Dynamic Surround Affects Response Amplitude and Kinetics.



S2 Fig. Dynamic Surround and Input Amplitude Do Not Affect L1 Response Kinetics.

3. DISCUSSION

3.1 Contrast Normalization in the Fly Optic Lobe

In Manuscript 1 (Drews et al., 2020) and in Manuscript 2 (Pirogova and Borst, under review), we discovered that the canonic computation of contrast normalization is present in the early visual system of *Drosophila*, comprehensively described its characteristics and its effect on the response of the cells, and demonstrated the importance of contrast normalization for motion vision.

3.1.1 Effects of Contrast Normalization

In Manuscript 1, we demonstrated contrast normalization in the medulla cells with band-pass filtering properties (Mi1, Tm3, Tm1, and Tm2) that serve as their main inputs into the first direction-selective neurons in the *Drosophila* optic lobe (T4 and T5). We showed that a dynamic visual surround suppresses the amplitude of the cells' response (Manuscript 1) and speeds it up, making the signal more transient (Manuscript 2). Both of these findings are similar to results that had been obtained from neurons in the vertebrate retina (Chander and Chichilinsky, 2001; Kim and Rieke, 2001; Baccus and Meister, 2002). We also showed that different cell types are affected by contrast normalization to a different extent.

3.1.2 Characteristics of Contrast Normalization

In Manuscript 1, we demonstrated that contrast normalization arises early in the optic lobe of the fly: while not present in the lamina, the first neuropil after the retina, contrast normalization already affects the cells one synapse downstream, in the medulla, both in the ON and the OFF visual pathways. From the medulla, the effects of contrast normalization propagate further to the first direction-selective cells.

In Manuscript 1, we showed that contrast normalization in response to visual surround motion is fast (<50 ms) and does not depend on the direction of the motion, with all directions suppressing the response amplitude to the same degree. Moreover, in Manuscript 1 and Manuscript 2, we determined that contrast normalization does not require a spatial component and that the normalizing effect on the signal is caused by dynamic but not by stationary stimuli in the visual surround. For example, we showed that a low-contrast moving grating exerted a suppressive effect on the amplitude, while a high-contrast grating that was

stationary or moving at a frequency beyond that discernible by the fly's eye did not. A high-contrast surround that was dynamic but had no motion component also exerted a suppressive effect on the amplitude of the response.

We demonstrated that the effects of contrast normalization come from outside the receptive field of the cell. In Manuscript 2, we confirmed that the visual surround beyond the cell's receptive field affects the cell's response to a stimulus covering a small central portion of its receptive field. And in Manuscript 1, we identified that the normalizing pool is integrated over a large spatial scale, with the strength of the normalization growing as the size of the visual surround increases.

3.1.3 Contrast Normalization Improves Robustness of Motion Detectors

The existing models of motion detection in *Drosophila* are often used to replicate the responses of large-field LPTC neurons. When presented with natural stimuli, however, the models underperform compared to the electrophysiological recordings from the cells. While the models successfully match the direction of stimulus motion, they, in contrast to LPTCs, fail to consistently match the velocity of the stimulus.

In Manuscript 1, we confirmed that adding a compression stage, equivalent to the task of contrast normalization, to the existing model dramatically improves its performance, putting it on par with the LPTC responses. Consequently, we determined the importance of contrast normalization for motion detection and for robust responses to natural images.

3.2 Mechanism of Contrast Normalization

In addition to establishing the presence and characterizing the main properties of contrast normalization in the early visual system of *Drosophila*, we also investigated its mechanism. We demonstrated that contrast normalization is implemented via a feedback loop, proved that the effects on the amplitude and temporal dynamics of the response are not a result of signal saturation, emphasized the potential role of shunting inhibition in the normalization process, and related the effects of normalization to a known mechanism in the fly's optic lobe that had also been shown to increase the temporal dynamics of cell responses.

3.2.1 Contrast Normalization Relies on Neural Feedback

In Manuscript 1, we proposed two alternative models to explain the origin of the normalizing signal in the medulla neurons (Mi1, Tm3, Tm1, and Tm2): 1.) A feedforward

model, in which the output of the upstream neuropil, the lamina, constitutes the normalizing pool for the medulla cells, or 2.) a feedback model, in which the signal from the cell itself is fed back to create the normalizing pool.

To distinguish between the two models, we disrupted the feedback connections from the cells of interest by blocking their synaptic output and measured the extent of contrast normalization. The absence of synaptic feedback dramatically reduced the suppressive effect of contrast normalization on the response amplitude of the cell of interest. The degree of this reduction varied between cell types but, even when the feedback was lacking, the effects of contrast normalization were still present for all cell types. Thus, the feedback signal is critical for contrast normalization but is not its sole component.

Although the significant reduction but not a complete absence of contrast normalization might be a result of an incomplete genetic block, it is more likely a consequence of the complex architecture of the feedback circuit. As such, the normalizing pool is likely to encompass more than just the output signal of one cell type alone, and contributions from other cell types might be sufficient for contrast normalization to take place to a lesser degree. Alternatively, with the feedback from the cell lacking, the weights given to other inputs into the normalizing pool might be altered, allowing contributions from other cells to compensate for this absence.

Since contrast normalization relies on neuronal feedback, the next step should be to identify the exact elements that constitute this feedback circuit.

3.2.2 Contrast Normalization Alters Response Time Constants

The effects of contrast normalization can be best represented by incorporating a divisive nonlinearity, either static or dynamic. A static nonlinearity is well described by a saturating contrast dependency, in which different conditions of the stimulus surround are represented by varying the saturation constant. Here, a high-contrast isolated local stimulus causes the response to saturate, further leading to changes in the response amplitude and temporal dynamics. Alternatively, a dynamic nonlinearity takes into account the biophysical properties of the neuronal membrane and entails a change in the filtering properties of the cell, i.e., in the membrane time constant.

In Manuscript 2, we confirmed that the model with a static nonlinearity does not faithfully replicate the effects of contrast normalization seen in the responses of the medulla interneurons. We, thus, concluded that the increased speed of the temporal dynamics and the suppressed amplitude of the signal at higher contrasts are not consequences of a simple

response saturation. The mechanism of contrast normalization, therefore, involves a change in the filtering properties of the cell, specifically entailing a change in its time constant. As the time constant is a function of the membrane resistance, this change in the resistance of the cell membrane can help elucidate the role of shunting inhibition in the process of contrast normalization.

3.2.3 Contrast Normalization and Shunting Inhibition

Shunting inhibition inhibits postsynaptic potential by locally reducing the input resistance (Figure 6). In contrast to other types of inhibition, such as hyperpolarization, which effects are best represented by linear subtraction, shunting inhibition most closely corresponds to the mathematical operation of division.

Shunting inhibition had previously been incorporated into contrast normalization models, such as in the model from Carandini and Heeger (1994) and Carandini et al. (1997):

$$g = \frac{g_0}{\sqrt{1-k\Sigma R}},$$

where, g corresponds to the conductance, ΣR to the pooled neuronal activity that normalizes the signal, and the parameter k determines the effectiveness of the normalization pool.

Here, the authors extended a simple linear model to include mutual shunting inhibition among a large number of cells to explain normalization in simple cells of the macaque primary visual cortex. In the model, shunting inhibition normalized the linear response of the cell by the pooled activity from the signals of the surrounding neurons, with the strength of shunting inhibition increasing at higher stimulus contrasts, i.e., with the rise in the activity of the neuronal pool. The increase in shunting inhibition also caused an increase in the conductance of the cell, thus leading to a reduction of the cell's membrane time constant and a decrease in its integration time. Subsequently, the change in the membrane time constant altered the temporal dynamics of the response in a manner similar to the effects of contrast normalization found in vertebrates (Chander and Chichilinsky, 2001; Kim and Rieke, 2001; Baccus and Meister, 2002) and in Manuscript 2, namely, by increasing the speed of the responses at higher stimulus contrasts. Overall, the normalization model that incorporates

shunting inhibition, as proposed by Carandini et al. (1997), successfully accounts for the response nonlinearities.

As shunting inhibition corresponds to the change in the input resistance of the cell, comparing the membrane resistance at different stimulus contrasts might of further interest.

3.2.4 State- and Stimulus-Dependent Modulation of Neural Processing Speed

The increase in the speed of the medulla interneuron responses at higher contrasts of the visual surround is reminiscent of the change that octopamine induces in the signals of *Drosophila* visual neurons (Suver et al., 2012).

Octopamine is a neurotransmitter in the optic lobe of *Drosophila* that plays an important role in flight (Brembs et al., 2007; Jung et al., 2011) and is involved in the state-dependent modulation of neural processing speed. Namely, during active locomotion, *Drosophila* visual neurons exhibit a boost in their response when compared to a resting state (Chiappe et al., 2010; Maimon et al., 2012; Suver et al., 2012), and the endogenous release of octopamine is involved in this state-dependent modulation (Suver et al., 2012). Moreover, this boost in the cell temporal filtering can be mimicked through the application of octopamine (Suver et al., 2012) and octopamine receptor agonist chlordimeform (CDM) (Arenz et al., 2017).

In active locomotion, the fly is likely to be confronted with visual inputs that are more dynamic than the inputs encountered when the fly is at rest, with a significant portion of the dynamics induced by the fly's self-motion. Thus, it is easy to draw parallels between the resting and active locomotion states of the fly and the static and dynamic categories of visual surround stimuli in Manuscript 2 and, thus, the role of visual surround in contrast normalization and temporal filtering of the signal. Moreover, visual stimuli alone, without any change in the activity of the fly, have been shown to elicit changes in the shape of the response similar to those produced by octopamine (Kohn et al., 2021). Therefore, the effect observed in Manuscript 2 might be of a similar nature, reflecting the activity-dependent modulations in the responses of a sensory system. Specifically, the modulation shifts the neuronal responses to higher frequencies, which the fly is more likely to encounter in locomotion, altering the processing characteristics to fit a specific behavioral state.

Recapitulating the effect of high-contrast surround stimuli using octopamine and studying the similarities between the two mechanisms, then, is of further interest.

3.3 Implications for Characterization of Visual Neurons

The effect of contrast normalization on the filtering properties of the visual neurons has serious implications for characterizing single cells and building motion detector models. Knowing the impact that the stimulus design has on the neuronal responses, however, might help us understand the observed inconsistencies in cells' filtering properties, observed in different studies. It might also offer a solution: using a consistent stimulus throughout the motion vision circuit to acquire a comprehensive description of cells' characteristics for the motion detector models.

3.3.1 Stimuli to Characterize Filtering Properties of the Cell

Characterizing the filtering properties of single cells creates an understanding of individual elements of a complex circuit, which is an important step that needs to be carried out to faithfully replicate biological mechanisms algorithmically. Current genetic tools make it possible to isolate specific neuronal populations as well as single cells for access and manipulation and to observe their responses in a tightly controlled environment.

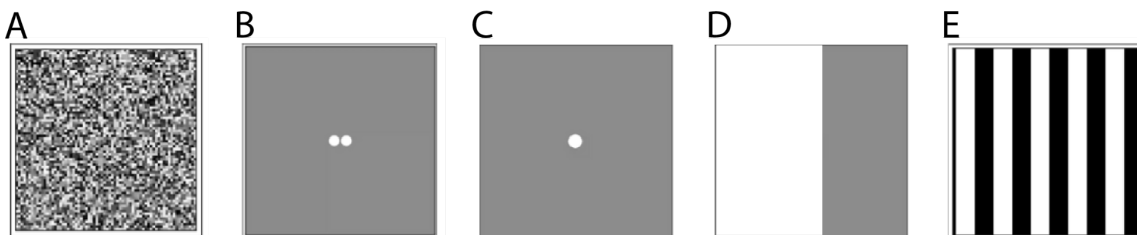


Figure 7. Stimuli to characterize cell filtering properties. Examples of A. stochastic noise, B. apparent motion, C. light flickers, and D. moving edges and E. gratings.

A plethora of artificial stimuli used to interrogate the visual system includes stochastic noise, light flickers, apparent motion, and moving edges and gratings, to name only a few (Figure 7). One of the reasons for such a variety is the increase in feature selectivity and in the complexity of computations exhibited by the cells when going down the visual pathway (discussed in Introduction 4.1 Computations in the *Drosophila* Optic Lobe). While the same stimuli are sometimes used to stimulate neurons in different neuropils, often stimuli that are specifically tailored to the cells of interest elicit the best response and make it possible to study the exact aspects of a mechanism. For example, light flickers are commonly used to determine the preferred polarity of the cells in the lamina (Clark et al., 2011). In the medulla, stochastic stimuli that allow for the extraction of the spatiotemporal RF are used to find cells'

receptive fields and characterize their impulse responses (Arenz et al., 2017). In the lobula and lobula plate, moving edges and gratings are often used to determine the preferred direction of motion of the first direction-selective cells and describe the properties of the motion detector (Serbe et al., 2016; Maisak et al., 2013). Moreover, apparent motion is used to stimulate the first direction-selective cells to determine if their response to motion is more or less than the sum of their responses to distinct steps that constitute the motion (Haag et al., 2016).

Each of these synthetic stimuli offers numerous advantages for characterizing cell responses: all of the main stimulus parameters (mean luminance, contrast, velocity, etc.) are fully controlled by the experimenter, and the stimuli themselves are easy to tailor and cater to the specific preferences of the cell (the cell's preferred polarity, its receptive field size, etc.). If the cells in question are linear filters, then any of these stimuli can fully characterize their responses. However, as the cells of interest often exhibit non-linear filtering properties, these synthetic stimuli might potentially create obstacles when response properties characterized with them are used to predict cell responses to naturalistic images. Unlike natural images, many of the synthetic stimuli are locally focused to stimulate single cells, neglecting the interactions of a neural network, and their controlled parameters lack the richness of cues found in complex natural environments. Thus, it would not be surprising if differences in stimulus features have an impact on the cell response properties

A major difference between the features of the synthetic stimuli used in the *Drosophila* optic lobe is their spatial profile, i.e., the presence or absence of the visual surround and its dynamics. For example, two large groups of stimuli can be distinguished: local (focused light flickers, apparent motion, etc.) and global (full-field gratings and edges, stochastic noise stimuli, etc.). However, as we have shown in Manuscript 1 and Manuscript 2, the presence or absence of dynamic stimuli in the visual surround, far outside of the cell's receptive field, has a pronounced influence on the response amplitude as well as on the temporal response properties of the cell. These findings exemplify the impact that the differences in stimulus design have on the cell responses and their interpretation.

3.3.2 Stimulus Design and Inconsistencies in Cell Characterization

Even small variations in stimulus design might cause divergent responses from the same cell, leading us to interpret its filtering properties differently. For example, in Manuscript 2, we saw that the sensitivity and the temporal filtering properties of the medulla cells that exhibit contrast normalization vary dramatically depending on the visual surround. For

instance, Mi1, in the absence of the visual surround, responded prominently to even the lowest amplitude of the visual input and, at first glance, also possessed temporal dynamics reminiscent of a low-pass filter, responding with a sustained signal throughout the presentation of the visual stimulus. This effect was, however, reversed in the presence of a high-contrast dynamic surround, which made the cell significantly less sensitive to the changes in the input stimulus but also altered the temporal dynamics of the response to that of a band-pass filter, with the response of the cell already starting to decay during the stimulus presentation. This divergence in the interpretation of dissimilar cell responses, caused by different visual surround conditions, potentially holds true for other visual neurons that exhibit contrast normalization.

The impact that differences in stimulus design have when studying similar phenomena is evident when comparing Manuscript 1 and Manuscript 2 to a recent study of temporal contrast adaptation in the early *Drosophila* visual system by Matulis et al. (2020). Matulis et al. (2020) also used two-photon calcium imaging to scan the main inputs into the motion vision circuit, probing a number of neurons that were also investigated in Manuscript 1 and Manuscript 2. Matulis et al. (2020), however, used full-field, binary noise that switched between periods of low and high contrast, with the change in stimulus contrast simultaneously applying to both the cells' RF and its visual surround (see below). Similarly to Manuscript 1 and Manuscript 2, Matulis et al. (2020) found contrast normalization to be present in the neurons in the motion vision circuit, both in the ON and the OFF pathway, and showed that the models of motion detection performed better at estimating the velocity of moving naturalistic images when they included a contrast adaptation stage.

However, the majority of the cells that Matulis et al. (2020) found to be affected by contrast normalization are distinct from the ones identified in Manuscript 1. And, conversely, many neurons that, in Manuscript 1, exhibited contrast normalization, were not identified as such by Matulis et al. (2020). Additionally, the few neurons that were shown to exhibit contrast normalization in both studies, varied in the degree to which they were affected by the visual contrast. Furthermore, Matulis et al. (2020) found that contrast adaptation had no effect on the temporal dynamics of the response, contrary to the findings in Manuscript 2 and the effects of contrast normalization that had previously been described in the vertebrate retina (Chander and Chichilinsky, 2001; Kim and Rieke, 2001; Baccus and Meister, 2002).

These stark discrepancies may partly be attributed to the differences in the visual stimuli used to interrogate the motion vision circuit (Figure 8). Matulis et al. (2020) used a full-field stimulus that did not separate between the visual center and surround, i.e., the stimuli within

the cell's receptive field that directly stimulate the neuron of interest and the stimuli outside of the cell's receptive field that do not stimulate the neuron of interest when presented in isolation. It provided constant stimulation to the entirety of the cell's receptive field and beyond and potentially masked the effects of any inhibition coming from outside of the receptive field. Our stimuli, on the other hand, were designed to deliver the stimulation to the cell's receptive field and to its surround independently (Manuscript 1 and Manuscript 2) and then were further focused on stimulating only the central part of the cell's receptive field while providing the normalizing signal in the visual surround (Manuscript 2).

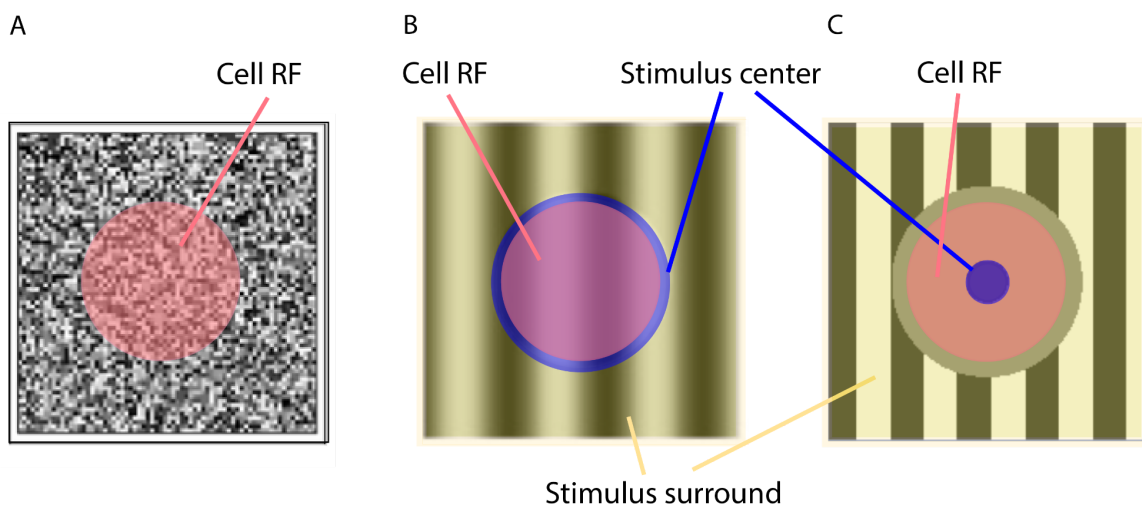


Figure 8. Stimuli spatial profiles. Examples of stimuli spatial profiles from A. Matulis et al. (2020), B. Manuscript 1, and C. Manuscript 2, showing the relation of the cell's receptive field (cell RF) to the stimulus center and surround.

This example illustrates the dramatic effects that differences in the stimuli, designed to study similar phenomena, can have on the cell responses and the interpretation of the results obtained from the same cells.

3.3.3 Stimuli Profiles and Motion Detection

The most commonly employed models of motion detection in the fly optic lobe are coincidence detectors (Figure 9), namely, Hassenstein-Reichardt (Hassenstein and Reichardt, 1956), Barlow-Levick (Barlow and Levick, 1965), and, more recently, Hassenstein-Reichardt/Barlow-Levick (Haag et al., 2016; Arenz et al., 2017), described in more detail in Introduction 4.2.2 Coincidence Detectors. In these models, the robustness of motion detection relies on the model inputs, the balance between their contributions, and their temporal filtering properties, which together allow the coincidence to occur and the

non-linear computation to produce reliable and reproducible results. As such, the effect that the surround of the stimulus on the filtering properties of the cell has a strong impact not only on the characterization of the visual interneurons themselves but also on the process of motion detection and the interpretation of the motion detector output.

Firstly, contrast normalization might affect the composition of motion detector models or, in general, which motion detector models are considered. When choosing cells as the motion detector inputs, we take into account their spatial characteristics and filtering properties, for example, we distinguish between slower low-pass and faster band-pass filter cells and select a combination thereof that would facilitate the coincidence in the detector. However, if the properties of the cell that are crucial for the motion detector models, such as time constants, vary as a result of contrast normalization, the measured cell properties will also differ depending on the stimulus used for their characterization. Thus, a different combination of input cells might be chosen to build a motion detector or a motion detector model might be favored with a different non-linear computation altogether. This way, different stimuli used to characterize cells can partly explain the divergence in the models proposed to algorithmically represent motion detection in the *Drosophila* optic lobe.

Secondly, contrast normalization might change the balance between the motion detector inputs. In Manuscript 2, we saw that contrast normalization does not alter by an equal amount the amplitude and the temporal dynamics of all the cells that it affects. Thus, the relative contribution of the different inputs to the motion-selective T4 and T5 cells will vary depending on the global structure of the stimulus. For example, if we look specifically at the ON pathway, the main two inputs to direction-selective T4 cells are Mi1 and Tm3 neurons (Takemura et al., 2013). In Manuscript 2, we showed that, at higher contrasts of the surround, Tm3 responses are suppressed and sped up to a higher degree than the responses of Mi1 neurons. Thus, the balance between the contribution of Mi1 and Tm3 is different depending on the visual stimulus, for instance, under the conditions of global dynamic visual surround or when presented with a local stimulus.

Finally, the change in the time course of the inputs has potential implications for the temporal dynamics of the motion detector itself. By adjusting the inputs' velocity tuning to the prevailing speed in the surround, contrast normalization would also change the preferred frequency of T4 and T5 neurons downstream, i.e., alter the preferred frequency of the motion detector.

To faithfully replicate the motion detector in the models, the filtering properties of the neurons need to not be conflated with the influence of the visual surround. To this end, the whole motion vision circuit needs to be scanned with a consistent set of stimuli.

3.3.4 Synthetic Stimuli to Scan the Motion Detection Circuit

Even though using the same set of stimuli to scan the motion vision circuit, from the lamina inputs through the medulla interneurons to the motion detector itself, would not make it possible to generalize the results for all visual stimuli, it would provide a complete set of answers accurately characterizing motion detector's response to the particular set of stimulus parameters. This puts a lot of importance on finding the correct stimulus. The search is complicated by the fact that the cells along the motion vision pathway vary greatly in regard to their preferred stimuli. For example, pixel noise is commonly used to characterize spatiotemporal RFs of lamina and medulla cells but it elicits limited responses from the motion-sensitive T4 and T5 cells.

To accommodate for the cells' preferences while ensuring control over specific stimulus parameters and the stimulus delivery (e.g., being able to focus the stimulus on a single cell), we might employ synthetic stimuli with a simpler local architecture and a motion component or an illusion of motion, such as apparent motion. However, using local synthetic stimuli to probe neurons that respond to more complex stimulus features raises the question if the stimuli themselves might cause a change in the filtering properties of the cell as it is trying to accommodate for their artificial nature. For example, apparent motion is a stimulus that successfully stimulates cells in isolation, creating an illusion of motion that evokes responses throughout the motion vision circuit. It is, however, localized to a couple of bright points on the screen, although the cells have evolved in the presence of complex natural images and developed as a part of a larger network, the connections of which would remain silent due to the lack of inputs that a cell would receive naturally.

On the other hand, using natural images for stimulation will let us faithfully represent the environment in which the eye has evolved. At the same time, it might make it impossible to interpret the properties of the cell due to the lack of control over precise stimulus parameters and delivery, as well as the difficulty of disentangling the diversity of visual cues that the neurons are responding to.

Alternatively, global synthetic stimuli, such as full-field gratings or edges, might offer a middle ground between simplified synthetic stimuli and natural images to probe the motion vision circuit. Such full-field stimuli stimulate the whole neural network while still offering a

high level of control over their design. However, they do not provide access to the filtering properties of separate cells, potentially stimulating competing mechanisms that can mask each other's effects. For example, when a cell that exhibits contrast normalization is stimulated by a full-field high-contrast dynamic pattern, its response would be a combination of the suppressive effects of the contrast in the visual surround and the boost in the response due to the stimulation of the cell's RF itself, emphasizing the importance of separating visual stimulus center and surround to disentangle the effects of the normalization (such as in Manuscript 1 and Manuscript 2).

Seeing how the differences in the visual surround outside of the cell's RF affect the filtering properties of the cell raises the question: what is the right balance between simplifying the stimulus and staying true to the natural environments? And to what extent are filtering properties of the cells fixed or able to change to accommodate for the artificial nature of the stimulus?

Finding the ideal synthetic stimulus to probe the motion detection circuit might not be an easy task, but measuring the whole circuit under identical stimulus conditions would provide us with faithful signals for our models. A solution to this problem was offered in a recent study on T4 cells (Groschner et al., 2022) that scanned the motion vision circuit with a set of dynamic global stimuli at varying frequencies to determine the influence that visual surround has on temporal filtering properties and velocity tuning of the neurons.

3.4 Conclusion and Future Directions

We demonstrated that contrast normalization is present in the early visual system of *Drosophila*, confirmed its importance for the computations in the motion vision circuit, and further described some of the main characteristics of the normalization process. Additionally, we narrowed down the space of possible mechanisms that explain how contrast adaptation shapes motion detection. Namely, we determined that the normalization is implemented via a feedback loop and suggested the role of shunting inhibition in the normalization process. The future investigation of the normalization mechanism should focus on describing further characteristics of the contrast normalization process, as well as searching for the main players in the normalizing circuit. Below, some of the future avenues to explore are described.

3.4.1 Searching for the Normalizing Neuron

In Manuscript 1, we confirmed that contrast normalization relies on a feedback connection, however, blocking the output of the neuron that exhibits normalization did not completely abolish the normalization effects in that neuron. Additionally, as the neurons exhibiting contrast normalization are excitatory, there must exist an inhibitory neuron, over which the normalization is carried out. Thus, another “normalizing neuron” plays a major role in this feedback loop, and, thus, the search for the normalizing neuron continues.

Even though we did not identify the normalizing neuron itself, we can describe some of its characteristics: 1.) The normalizing neuron is a temporal band-pass filter that responds transiently to the changes in the visual stimulus. As we have shown in Manuscript 2, the normalizing signal is present when the visual surround is dynamic but not when it is static, ruling out neurons with low-pass filtering characteristics that would respond in a sustained manner to a static surround; 2.) The normalizing neuron is not selective for the direction of motion (as we have shown in Manuscript 1) or for motion in general (as we have shown in Manuscript 2), as a visual surround that is temporally dynamic but has no motion component is enough to elicit contrast normalization; 3.) The normalizing neuron either possesses a large receptive field or forms an interconnected network within its cell type. As we have shown in Manuscript 2, the normalization signal comes from outside the cell's receptive field, and the strength of contrast normalization continues increasing beyond 85° from the receptive field center (Manuscript 1). 4.) The normalizing neuron provides inhibitory signals.

The next steps in the search for the normalizing neuron would involve tracing synaptic connections from the medulla neurons that exhibit contrast normalization (for example, using the connectome of the *Drosophila* optic lobe) and further focusing on the neurons that can integrate signals from multiple neurons into a normalizing pool, either by being a part of a large interconnected network themselves or by possessing dense arborizations that span several columns. After the candidates are identified, the normalization circuit can be manipulated by blocking the output of the candidate neurons and measuring the impact of the block on the normalization effects. Alternatively, the search can be narrowed down by characterizing the filtering properties and receptive fields of the candidate neurons to find the best match to the characteristics described above. Another way to narrow down the search is by looking at the neurotransmitters of the candidate neurons and focusing on the ones that release either GABA or glutamate, for which receptors in the fly nervous system are mostly inhibitory (Liu and Wilson, 2013; Mauss et al., 2014; Davis et al., 2018).

A potential obstacle to a straightforward interpretation of the results might emerge if the normalizing neuron is, in fact, itself a part of a larger network that can partially compensate for the lack of input from a cell type, masking the effects of the block.

3.4.2 Validating the Role of Shunting Inhibition

In Manuscript 2, we proposed a model that incorporates a dynamic nonlinearity to explain the mechanism of contrast normalization and confirmed that the model was in agreement with our experimental results. From the model, we hypothesized that the change in membrane resistance is involved in contrast normalization, thus implying that shunting inhibition is likely to play an important role in the mechanism via which the normalization is implemented. We further postulated that the input resistance of the cell should drop significantly in the presence of a dynamic visual surround.

Measuring the resistance of the cell membrane under different contrast conditions would allow us to test this hypothesis, confirming or denying the role of shunting inhibition and further characterizing properties of contrast normalization.

As it is not possible to carry out electrophysiological recordings from the dendrites of the medulla interneurons, the measurements would need to be taken at the soma. However, for compartmentalized neurons, it might not be possible to detect the drop in the input resistance, that occurs at the dendrites, from the soma. This would delay drawing unambiguous conclusions and validating our hypothesis until the technology is available to perform the local measurements at the cell's dendrites.

3.4.3 Recapitulating the Effects of the Dynamic Surround with Octopamine

Octopamine speeds up the responses of *Drosophila* visual interneurons when the fly is in active locomotion (Suver et al., 2012), similar to the effect of contrast normalization on the cell responses that we observed in Manuscript 2 in the presence of a dynamic visual surround. Moreover, when in flight, the visual inputs that the fly receives are more dynamic than when the fly is at rest, with the visual environment it encounters potentially reflecting the difference between the static and dynamic visual surround used in Manuscript 2.

That is why the next step is to explore a potential connection between the drop in the input resistance and the activation of octopamine receptors (specifically, Oamb and Oct1 β R octopamine receptors (El-Kholy et al., 2015) have been shown to be expressed in the *Drosophila* optic lobe). This can be done via a pharmacological application of octopamine or CDM and would allow us to shed further light on the mechanism of contrast normalization.

4. REFERENCES

- Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, 275(5297), 221–224. <https://doi.org/10.1126/science.275.5297.221>
- Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F., George, R. A., Lewis, S. E., Richards, S., Ashburner, M., Henderson, S. N., Sutton, G. G., Wortman, J. R., Yandell, M. D., Zhang, Q., ... Venter, J. C. (2000). The genome sequence of *Drosophila melanogaster*. *Science*, 287(5461), 2185–2195. <https://doi.org/10.1126/science.287.5461.2185>
- Arenz, A., Drews, M. S., Richter, F. G., Ammer, G., & Borst, A. (2017). The temporal tuning of the *Drosophila* motion detectors is determined by the dynamics of their input elements. *Current Biology*, 27(7), 929–944. <https://doi.org/10.1016/j.cub.2017.01.051>
- Baccus, S. A., & Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36(5), 909–919. [https://doi.org/10.1016/S0896-6273\(02\)01050-4](https://doi.org/10.1016/S0896-6273(02)01050-4)
- Barlow, H. B., & Rosenblith, W. A. (1961). Possible principles underlying the transformations of sensory messages (pp. 217–234). *MIT Press*.
- Barlow, H. B., & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *The Journal of Physiology*, 178(3), 477–504. <https://doi.org/10.1113/jphysiol.1965.sp007638>
- Baylor, D. A., & Fuortes, M. G. F. (1970). Electrical responses of single cones in the retina of the turtle. *The Journal of Physiology*, 207(1), 77–92. <https://doi.org/10.1113/jphysiol.1970.sp009049>
- Borst, A., & Egelhaaf, M. (1989). Principles of visual motion detection. *Trends in Neurosciences*, 12(8), 297–306. [https://doi.org/10.1016/0166-2236\(89\)90010-6](https://doi.org/10.1016/0166-2236(89)90010-6)
- Borst, A., & Abarbanel, H. di. (2007). Relating a calcium indicator signal to the unperturbed calcium concentration time-course. *Theoretical Biology and Medical Modelling*, 4(1), 7. <https://doi.org/10.1186/1742-4682-4-7>
- Borst, A., Drews, M., & Meier, M. (2020). The neural network behind the eyes of a fly. *Current Opinion in Physiology*, 16, 33–42. <https://doi.org/10.1016/j.cophys.2020.05.004>
- Borst, A., Flanagan, V. L., & Sompolinsky, H. (2005). Adaptation without parameter change: Dynamic gain control in motion detection. *Proceedings of the National Academy of Sciences*, 102(17), 6172–6176. <https://doi.org/10.1073/pnas.0500491102>

- Brand, A. H., & Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, *118*(2), 401–415. <https://doi.org/10.1242/dev.118.2.401>
- Brembs, B., Christiansen, F., Pfluger, H. J., & Duch, C. (2007). Flight initiation and maintenance deficits in flies with genetically altered biogenic amine levels. *Journal of Neuroscience*, *27*(41), 11122–11131. <https://doi.org/10.1523/JNEUROSCI.2704-07.2007>
- Brown, S. P., & Masland, R. H. (2001). Spatial scale and cellular substrate of contrast adaptation by retinal ganglion cells. *Nature Neuroscience*, *4*(1), 44–51. <https://doi.org/10.1038/82888>
- Cai, L. T., Krishna, V., Hladnik, T. C., Guilbeault, N. C., Juntti, S. A., Thiele, T. R., Arrenberg, A. B., & Cooper, E. A. (2020). Visual statistics of aquatic environments in the natural habitats of zebrafish. *Journal of Vision*, *20*(11), 433. <https://doi.org/10.1167/jov.20.11.433>
- Camp, A. J., Cheong, S. K., Tailby, C., & Solomon, S. G. (2011). The impact of brief exposure to high contrast on the contrast response of neurons in primate lateral geniculate nucleus. *Journal of Neurophysiology*, *106*(3), 1310–1321. <https://doi.org/10.1152/jn.00943.2010>
- Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science*, *264*(5163), 1333–1336. <https://doi.org/10.1126/science.8191289>
- Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *The Journal of Neuroscience*, *17*(21), 8621–8644. <https://doi.org/10.1523/JNEUROSCI.17-21-08621.1997>
- 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>
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science*, *263*(5148), 802–805. <https://doi.org/10.1126/science.8303295>
- Chamberland, S., Yang, H. H., Pan, M. M., Evans, S. W., Guan, S., Chavarha, M., Yang, Y., Salesse, C., Wu, H., Wu, J. C., Clandinin, T. R., Toth, K., Lin, M. Z., & St-Pierre, F. (2017). Fast two-photon imaging of subcellular voltage dynamics in neuronal tissue with genetically encoded indicators. *ELife*, *6*. <https://doi.org/10.7554/eLife.25690>
- Chander, D., & Chichilnisky, E. J. (2001). Adaptation to temporal contrast in primate and salamander retina. *The Journal of Neuroscience*, *21*(24), 9904–9916. <https://doi.org/10.1523/JNEUROSCI.21-24-09904.2001>

- Chen, T.-W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K., & Kim, D. S. (2013). Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, *499*(7458), 295–300. <https://doi.org/10.1038/nature12354>
- Chiappe, M. E., Seelig, J. D., Reiser, M. B., & Jayaraman, V. (2010). Walking modulates speed sensitivity in *Drosophila* motion vision. *Current Biology*, *20*(16), 1470–1475. <https://doi.org/10.1016/j.cub.2010.06.072>
- Chichilnisky, E. J., & Kalmar, R. S. (2002). Functional asymmetries in ON and OFF ganglion cells of primate retina. *The Journal of Neuroscience*, *22*(7), 2737–2747. <https://doi.org/10.1523/JNEUROSCI.22-07-02737.2002>
- Chichilnisky, E. J. (2001). A simple white noise analysis of neuronal light responses. *Network (Bristol, England)*, *12*(2), 199–213.
- Clark, D. A., Bursztyn, L., Horowitz, M. A., Schnitzer, M. J., & Clandinin, T. R. (2011). Defining the computational structure of the motion detector in *Drosophila*. *Neuron*, *70*(6), 1165–1177. <https://doi.org/10.1016/j.neuron.2011.05.023>
- Clemens, J., Ozeri-Engelhard, N., & Murthy, M. (2018). Fast intensity adaptation enhances the encoding of sound in *Drosophila*. *Nature Communications*, *9*(1), 134. <https://doi.org/10.1038/s41467-017-02453-9>
- Coen-Cagli, R., & Solomon, S. S. (2019). Relating divisive normalization to neuronal response variability. *The Journal of Neuroscience*, *39*(37), 7344–7356. <https://doi.org/10.1523/JNEUROSCI.0126-19.2019>
- Cohen, L. (1989). Special topic: optical approaches to neuron function. *Annual Review of Physiology*, *51*(1), 487–490. <https://doi.org/10.1146/annurev.ph.51.030189.002415>
- Davis, F. P., Nern, A., Picard, S., Reiser, M. B., Rubin, G. M., Eddy, S. R., & Henry, G. L. (2020). A genetic, genomic, and computational resource for exploring neural circuit function. *ELife*, *9*, e50901. <https://doi.org/10.7554/eLife.50901>
- Demb, J. B. (2008). Functional circuitry of visual adaptation in the retina. *The Journal of Physiology*, *586*(18), 4377–4384. <https://doi.org/10.1113/jphysiol.2008.156638>
- Denk, W., Strickler, J. H., & Webb, W. W. (1990). Two-photon laser scanning fluorescence microscopy. *Science*, *248*(4951), 73–76. <https://doi.org/10.1126/science.2321027>
- Drews, M. S., Leonhardt, A., Pirogova, N., Richter, F. G., Schuetzenberger, A., Braun, L., Serbe, E., & Borst, A. (2020). Dynamic Signal Compression for Robust Motion Vision in Flies. *Current Biology*, *30*(2), 209–221.e8. <https://doi.org/10.1016/j.cub.2019.10.035>

- Eggermont, J. J., Johannesma, P. I. M., & Aertsen, A. M. H. J. (1983). Reverse-correlation methods in auditory research. *Quarterly Reviews of Biophysics*, *16*(3), 341–414. <https://doi.org/10.1017/S0033583500005126>
- El-Kholy, S., Stephano, F., Li, Y., Bhandari, A., Fink, C., & Roeder, T. (2015). Expression analysis of octopamine and tyramine receptors in *Drosophila*. *Cell and Tissue Research*, *361*(3), 669–684. <https://doi.org/10.1007/s00441-015-2137-4>
- Fischbach, K.-F., & Dittrich, A. P. M. (1989). The optic lobe of *Drosophila melanogaster*. I. A Golgi analysis of wild-type structure. *Cell and Tissue Research*, *258*(3), 441–475. <https://doi.org/10.1007/BF00218858>
- Fisher, Y. E., Silies, M., & Clandinin, T. R. (2015). Orientation selectivity sharpens motion detection in *Drosophila*. *Neuron*, *88*(2), 390–402. <https://doi.org/10.1016/j.neuron.2015.09.033>
- Frolov, R. v. (2016). Current advances in invertebrate vision: insights from patch-clamp studies of photoreceptors in apposition eyes. *Journal of Neurophysiology*, *116*(2), 709–723. <https://doi.org/10.1152/jn.00288.2016>
- Garvert, M. M., & Gollisch, T. (2013). Local and global contrast adaptation in retinal ganglion cells. *Neuron*, *77*(5), 915–928. <https://doi.org/10.1016/j.neuron.2012.12.030>
- Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, *59*(1), 167–192. <https://doi.org/10.1146/annurev.psych.58.110405.085632>
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*(1), 535–574. <https://doi.org/10.1146/annurev.neuro.29.051605.113038>
- Gonzalez-Bellido, P. T., Wardill, T. J., & Juusola, M. (2011). Compound eyes and retinal information processing in miniature dipteran species match their specific ecological demands. *Proceedings of the National Academy of Sciences*, *108*(10), 4224–4229. <https://doi.org/10.1073/pnas.1014438108>
- Götz, K. G. (1964). Optomotorische Untersuchung des visuellen systems einiger Augenmutanten der Fruchtfliege *Drosophila*. *Kybernetik*, *2*(2), 77–92. <https://doi.org/10.1007/BF00288561>
- Groschner, L. N., Malis, J. G., Zuidinga, B., & Borst, A. (2022). A biophysical account of multiplication by a single neuron. *Nature*, *603*(7899), 119–123. <https://doi.org/10.1038/s41586-022-04428-3>

- Gruntman, E., Romani, S., & Reiser, M. B. (2018). Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in *Drosophila*. *Nature Neuroscience*, *21*(2), 250–257. <https://doi.org/10.1038/s41593-017-0046-4>
- Haag, J., Arenz, A., Serbe, E., Gabbiani, F., & Borst, A. (2016). Complementary mechanisms create direction selectivity in the fly. *ELife*, *5*. <https://doi.org/10.7554/eLife.17421>
- Haag, J., Mishra, A., & Borst, A. (2017). A common directional tuning mechanism of *Drosophila* motion-sensing neurons in the ON and in the OFF pathway. *ELife*, *6*. <https://doi.org/10.7554/eLife.29044>
- Harmening, W. M. (2017). Kontrastempfindlichkeit und Sehschärfe bei Tieren. *Der Ophthalmologe*, *114*(11), 986–996. <https://doi.org/10.1007/s00347-017-0561-4>
- Harris, R. A., O'Carroll, D. C., & Laughlin, S. B. (2000). Contrast gain reduction in fly motion adaptation. *Neuron*, *28*(2), 595–606. [https://doi.org/10.1016/S0896-6273\(00\)00136-7](https://doi.org/10.1016/S0896-6273(00)00136-7)
- Hassenstein, B., & Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenbewertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift Für Naturforschung B*, *11*(9–10), 513–524. <https://doi.org/10.1515/znb-1956-9-1004>
- Hausen, K. (1976). Functional Characterization and Anatomical Identification of Motion Sensitive Neurons in the Lobula plate of the Blowfly *Calliphora erythrocephala*. *Zeitschrift Für Naturforschung C*, *31*(9–10), 629–634. <https://doi.org/10.1515/znc-1976-9-1001>
- Hausen, K. (1984). The lobula-complex of the fly: structure, function and significance in visual behaviour. In *Photoreception and Vision in Invertebrates* (pp. 523–559). Springer US. https://doi.org/10.1007/978-1-4613-2743-1_15
- Jarsky, T., Cembrowski, M., Logan, S. M., Kath, W. L., Rieke, H., Demb, J. B., & Singer, J. H. (2011). A synaptic mechanism for retinal adaptation to luminance and contrast. *Journal of Neuroscience*, *31*(30), 11003–11015. <https://doi.org/10.1523/JNEUROSCI.2631-11.2011>
- Jenett, A., Rubin, G. M., Ngo, T.-T. B., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B. D., Cavallaro, A., Hall, D., Jeter, J., Iyer, N., Fetter, D., Hausenfluck, J. H., Peng, H., Trautman, E. T., Svirskas, R. R., Myers, E. W., Iwinski, Z. R., Aso, Y., ... Zugates, C. T. (2012). A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Reports*, *2*(4), 991–1001. <https://doi.org/10.1016/j.celrep.2012.09.011>
- Joesch, M., Schnell, B., Raghu, S. V., Reiff, D. F., & Borst, A. (2010). ON and OFF pathways in *Drosophila* motion vision. *Nature*, *468*(7321), 300–304. <https://doi.org/10.1038/nature09545>

- Jung, S. N., Borst, A., & Haag, J. (2011). Flight activity alters velocity tuning of fly motion-sensitive neurons. *Journal of Neuroscience*, *31*(25), 9231–9237. <https://doi.org/10.1523/JNEUROSCI.1138-11.2011>
- Khani, M. H., & Gollisch, T. (2017). Diversity in spatial scope of contrast adaptation among mouse retinal ganglion cells. *Journal of Neurophysiology*, *118*(6), 3024–3043. <https://doi.org/10.1152/jn.00529.2017>
- Kim, K. J., & Rieke, F. (2001). Temporal contrast adaptation in the input and output signals of Salamander Retinal Ganglion Cells. *The Journal of Neuroscience*, *21*(1), 287–299. <https://doi.org/10.1523/JNEUROSCI.21-01-00287.2001>
- Kohn, J. R., Portes, J. P., Christenson, M. P., Abbott, L. F., & Behnia, R. (2021). Flexible filtering by neural inputs supports motion computation across states and stimuli. *Current Biology*, *31*(23), 5249–5260.e5. <https://doi.org/10.1016/j.cub.2021.09.061>
- Kolodziejczyk, A., Sun, X., Meinertzhagen, I. A., & Nässel, D. R. (2008). Glutamate, GABA and acetylcholine signaling components in the lamina of the *Drosophila* visual system. *PLoS ONE*, *3*(5), e2110. <https://doi.org/10.1371/journal.pone.0002110>
- Land, M. F. (1997). Visual acuity in insects. *Annual Review of Entomology*, *42*(1), 147–177. <https://doi.org/10.1146/annurev.ento.42.1.147>
- Laughlin, S. (1981). A simple coding procedure enhances a neuron's information capacity. *Zeitschrift Für Naturforschung C*, *36*(9–10), 910–912. <https://doi.org/10.1515/znc-1981-9-1040>
- Laughlin, S. B. (1989). The role of sensory adaptation in the retina. *Journal of Experimental Biology*, *146*(1), 39–62. <https://doi.org/10.1242/jeb.146.1.39>
- Leong, J. C. S., Esch, J. J., Poole, B., Ganguli, S., & Clandinin, T. R. (2016). Direction selectivity in *Drosophila* emerges from preferred-direction enhancement and null-direction suppression. *Journal of Neuroscience*, *36*(31), 8078–8092. <https://doi.org/10.1523/JNEUROSCI.1272-16.2016>
- Liu, J. K., & Gollisch, T. (2015). Spike-triggered covariance analysis reveals phenomenological diversity of contrast adaptation in the retina. *PLOS Computational Biology*, *11*(7), e1004425. <https://doi.org/10.1371/journal.pcbi.1004425>
- Liu, W. W., & Wilson, R. I. (2013). Glutamate is an inhibitory neurotransmitter in the *Drosophila* olfactory system. *Proceedings of the National Academy of Sciences*, *110*(25), 10294–10299. <https://doi.org/10.1073/pnas.1220560110>

- Maimon, G., Straw, A. D., & Dickinson, M. H. (2010). Active flight increases the gain of visual motion processing in *Drosophila*. *Nature Neuroscience*, *13*(3), 393–399. <https://doi.org/10.1038/nn.2492>
- Maisak, M. S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., Schilling, T., Bahl, A., Rubin, G. M., Nern, A., Dickson, B. J., Reiff, D. F., Hopp, E., & Borst, A. (2013). A directional tuning map of *Drosophila* elementary motion detectors. *Nature*, *500*(7461), 212–216. <https://doi.org/10.1038/nature12320>
- Makman, M. H. (1970). Adenyl cyclase of cultured mammalian cells: activation by catecholamines. *Science*, *170*(3965), 1421–1423. <http://www.jstor.org/stable/1730882>
- Matulis, C. A., Chen, J., Gonzalez-Suarez, A. D., Behnia, R., & Clark, D. A. (2020). Heterogeneous temporal contrast adaptation in *Drosophila* direction-selective circuits. *Current Biology*, *30*(2), 222–236.e6. <https://doi.org/10.1016/j.cub.2019.11.077>
- Mauss, A. S., Meier, M., Serbe, E., & Borst, A. (2014). Optogenetic and pharmacologic dissection of feedforward inhibition in *Drosophila* motion vision. *Journal of Neuroscience*, *34*(6), 2254–2263. <https://doi.org/10.1523/JNEUROSCI.3938-13.2014>
- Meier, M., Serbe, E., Maisak, M. S., Haag, J., Dickson, B. J., & Borst, A. (2014). Neural circuit components of the *Drosophila* OFF motion vision pathway. *Current Biology*, *24*(4), 385–392. <https://doi.org/10.1016/j.cub.2014.01.006>
- Meinertzhagen, I. A., & O’Neil, S. D. (1991). Synaptic organization of columnar elements in the lamina of the wild type in *Drosophila melanogaster*. *The Journal of Comparative Neurology*, *305*(2), 232–263. <https://doi.org/10.1002/cne.903050206>
- Nakai, J., Ohkura, M., & Imoto, K. (2001). A high signal-to-noise Ca²⁺ probe composed of a single green fluorescent protein. *Nature Biotechnology*, *19*(2), 137–141. <https://doi.org/10.1038/84397>
- Nikolaev, A., Leung, K.-M., Odermatt, B., & Lagnado, L. (2013). Synaptic mechanisms of adaptation and sensitization in the retina. *Nature Neuroscience*, *16*(7), 934–941. <https://doi.org/10.1038/nn.3408>
- Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2010). Divisive normalization in olfactory population codes. *Neuron*, *66*(2), 287–299. <https://doi.org/10.1016/j.neuron.2010.04.009>
- Osorio, D., & Vorobyev, M. (2005). Photoreceptor spectral sensitivities in terrestrial animals: adaptations for luminance and colour vision. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1574), 1745–1752. <https://doi.org/10.1098/rspb.2005.3156>
- Reichardt, W. (1987). Computation of optical motion by movement detectors. *Biophysical Chemistry*, *26*(2–3), 263–278. [https://doi.org/10.1016/0301-4622\(87\)80028-5](https://doi.org/10.1016/0301-4622(87)80028-5)

Rubin, G. M., & Spradling, A. C. (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science*, *218*(4570), 348–353. <https://doi.org/10.1126/science.6289436>

Ruff, D. A., Alberts, J. J., & Cohen, M. R. (2016). Relating normalization to neuronal populations across cortical areas. *Journal of Neurophysiology*, *116*(3), 1375–1386. <https://doi.org/10.1152/jn.00017.2016>

Sakmann, B., & Neher, E. (1984). Patch clamp techniques for studying ionic channels in excitable membranes. *Annual Review of Physiology*, *46*(1), 455–472. <https://doi.org/10.1146/annurev.ph.46.030184.002323>

Sanjayan, K. P. (2020). *Insect Physiology (21st Century Biology And Agriculture: Textbook Series)*.

Schnell, B., Raghu, S. V., Nern, A., & Borst, A. (2012). Columnar cells necessary for motion responses of wide-field visual interneurons in *Drosophila*. *Journal of Comparative Physiology A*, *198*(5), 389–395. <https://doi.org/10.1007/s00359-012-0716-3>

Serbe, E., Meier, M., Leonhardt, A., & Borst, A. (2016). Comprehensive characterization of the major presynaptic elements to the *Drosophila* OFF motion detector. *Neuron*, *89*(4), 829–841. <https://doi.org/10.1016/j.neuron.2016.01.006>

Shapley, R. M., & Victor, J. D. (1979). Nonlinear spatial summation and the contrast gain control of cat retinal ganglion cells. *The Journal of Physiology*, *290*(2), 141–161. <https://doi.org/10.1113/jphysiol.1979.sp012765>

Shinomiya, K., Karuppudurai, T., Lin, T.-Y., Lu, Z., Lee, C.-H., & Meinertzhagen, I. A. (2014). Candidate neural substrates for OFF-edge motion detection in *Drosophila*. *Current Biology*, *24*(10), 1062–1070. <https://doi.org/10.1016/j.cub.2014.03.051>

Shinomiya, K., Huang, G., Lu, Z., Parag, T., Xu, C. S., Aniceto, R., Ansari, N., Cheatham, N., Lauchie, S., Neace, E., Ogundeyi, O., Ordish, C., Peel, D., Shinomiya, A., Smith, C., Takemura, S., Talebi, I., Rivlin, P. K., Nern, A., ... Meinertzhagen, I. A. (2019). Comparisons between the ON- and OFF-edge motion pathways in the *Drosophila* brain. *eLife*, *8*. <https://doi.org/10.7554/eLife.40025>

Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology*, *13*(2), 144–149. [https://doi.org/10.1016/S0959-4388\(03\)00047-3](https://doi.org/10.1016/S0959-4388(03)00047-3)

Solomon, S. G., & Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. *Current Biology*, *24*(20), R1012–R1022. <https://doi.org/10.1016/j.cub.2014.09.001>

Suver, M. P., Mamiya, A., & Dickinson, M. H. (2012). Octopamine neurons mediate flight-induced modulation of visual processing in *Drosophila*. *Current Biology*, *22*(24), 2294–2302. <https://doi.org/10.1016/j.cub.2012.10.034>

Sweeney, S. T., Broadie, K., Keane, J., Niemann, H., & O’Kane, C. J. (1995). Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron*, *14*(2), 341–351. [https://doi.org/10.1016/0896-6273\(95\)90290-2](https://doi.org/10.1016/0896-6273(95)90290-2)

Takemura, S.-Y., Lu, Z., & Meinertzhagen, I. A. (2008). Synaptic circuits of the *Drosophila* optic lobe: The input terminals to the medulla. *The Journal of Comparative Neurology*, *509*(5), 493–513. <https://doi.org/10.1002/cne.21757>

Takemura, S., Nern, A., Chklovskii, D. B., Scheffer, L. K., Rubin, G. M., & Meinertzhagen, I. A. (2017). The comprehensive connectome of a neural substrate for ‘ON’ motion detection in *Drosophila*. *ELife*, *6*. <https://doi.org/10.7554/eLife.24394>

Tripp, B. P. (2012). Decorrelation of spiking variability and improved information transfer through feedforward divisive normalization. *Neural Computation*, *24*(4), 867–894. https://doi.org/10.1162/NECO_a_00255

Tuthill, J. C., Nern, A., Holtz, S. L., Rubin, G. M., & Reiser, M. B. (2013). Contributions of the 12 neuron classes in the fly lamina to motion vision. *Neuron*, *79*(1), 128–140. <https://doi.org/10.1016/j.neuron.2013.05.024>

Wark, B., Lundstrom, B. N., & Fairhall, A. (2007). Sensory adaptation. *Current Opinion in Neurobiology*, *17*(4), 423–429. <https://doi.org/10.1016/j.conb.2007.07.001>

Wilson, R. I., Turner, G. C., & Laurent, G. (2004). Transformation of olfactory representations in the *Drosophila* antennal lobe. *Science*, *303*(5656), 366–370. <https://doi.org/10.1126/science.1090782>

Wissig, S. C., & Kohn, A. (2012). The influence of surround suppression on adaptation effects in primary visual cortex. *Journal of Neurophysiology*, *107*(12), 3370–3384. <https://doi.org/10.1152/jn.00739.2011>

Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G. J., Tsang, A., Tsegaye, G., Narayan, S., Obara, C. J., Lim, J.-X., Patel, R., Zhang, R., Ahrens, M. B., Turner, G. C., Wang, S. S.-H., Korff, W. L., ... Looger, L. L. (2021). Fast and sensitive GCaMP calcium indicators for imaging neural populations. *BioRxiv*. <https://doi.org/10.1101/2021.11.08.467793>

Zheng, Z., Lauritzen, J. S., Perlman, E., Robinson, C. G., Nichols, M., Milkie, D., Torrens, O., Price, J., Fisher, C. B., Sharifi, N., Calle-Schuler, S. A., Kmecova, L., Ali, I. J., Karsh, B.,

Trautman, E. T., Bogovic, J. A., Hanslovsky, P., Jefferis, G. S. X. E., Kazhdan, M., ... Bock, D. D. (2018). A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*. *Cell*, *174*(3), 730-743.e22. <https://doi.org/10.1016/j.cell.2018.06.019>

5. APPENDIX

5.1 Acknowledgements

First and foremost, I would like to thank Axel for having given me the opportunity to be a part of his group and for being a great supervisor throughout the time of my PhD. I am grateful for the scientific and experimental freedom you gave me, allowing me to dive into the field of neuroscience. I thank you deeply for your trust and support.

I would like to thank my TAC members: Laura Busse, Christian Leibold, and Michael Myoga, – for their time and advice. I am grateful to the administrative team and members of the Graduate School for Systemic Neurosciences for their immeasurable support. I am deeply grateful to Lukas for proofreading my thesis and manuscript and for sparking numerous exciting scientific discussions. Thank you to Wolfgang, Stefan, Renate, Christian, and Romina for their technical assistance. Thank you to the whole Borst department, P1 and P7 especially, for sharing this experience with me.

Thank you to Michi for teaching me about two-photon imaging and modelling and for plentiful scientific discussions. And thank you to Sandra for teaching me most of what I know about the confocal and supporting me along the way. I am deeply grateful to Aicha for her incredible patience, her immense kindness, and the numerous movie nights! I am very thankful to Anna for supporting me in science and baking alike and for being there when I needed her most! A huge thank you to Maria who went on a similar bureaucratic journey and was there to talk about our adventures and our science! Finally, an enormous thanks goes to Leif and the cats who made even the hardest days bearable and the good days so much happier!

5.2 List of Manuscripts and Author Contributions

Manuscript 1

Drews, M. S.*, Leonhardt, A.*, **Pirogova, N.**, Richter, F. G., Schuetzenberger, A., Braun, L., Serbe, E., & Borst, A. (2020). Dynamic Signal Compression for Robust Motion Vision in Flies. *Current Biology*, 30(2), 209-221.e8.

*These authors contributed equally.

M.S.D., A.L., and A.B. jointly conceived the study. M.S.D. and A.L. designed all experiments. A.L. and L.B. conducted behavioral experiments. E.S. recorded electrophysiological responses. M.S.D., **N.P.**, F.G.R., and A.S. performed calcium imaging. A.L. designed and analyzed the convolutional model. M.S.D. and A.L. analyzed data, performed modeling, and wrote the manuscript. All authors participated in editing the manuscript.

I carried out experiments during the preliminary testing to assess cell responses and design the appropriate stimulus. For the final version, I carried out two-photon calcium imaging experiments of single cells of various cell types under a range of experimental conditions. The panels in Figure 3 and Figure 4 contain the data I contributed to this publication. Additionally, I commented on the manuscript

Drews, M. S.

Leonhardt, A..

Borst, A.

Manuscript 2

Pirogova, N., Borst, A.. Contrast Normalization Affects Response Time-Course of Visual Interneurons

N.P. and A.B. designed the study, **N.P.** conducted the experiments and analyzed the data, A.B. performed modeling, **N.P.** wrote the original draft, A.B. reviewed and edited the manuscript.

I planned the experiments and designed and programmed the stimulus. I conducted the two-photon calcium single-cell imaging for all the cell types and consequently wrote and carried out the data analysis. I wrote the original draft and further modified the manuscript.

Borst, A.

5.3 Curriculum Vitae

NADEZHDA PIROGOVA

PERSONAL PROFILE

Science communicator. Passionate about mental health advocacy and bringing scientific research closer to people. Experienced with team management, social media strategy development, and event coordination.

AREAS OF EXPERTISE

- ✓ Languages: English, Russian, German (B2 level)
- ✓ Programming: Python, MATLAB, R
- ✓ Applied computer skills: Adobe Illustrator, LaTeX, Hootsuite, Canva
- ✓ Science communication: short- and long-form copywriting, digital strategy development and implementation, content optimisation, team management, media relations, event planning and execution
- ✓ Certifications: Data Analyst Nanodegree (Udacity); Data Analyst, Data Scientist, and Python Programmer (DataCamp)

SCIENCE COMMUNICATION EXPERIENCE

Soapbox Science Munich

Social Media Manager and Event Coordinator | September 2019 - Present

- Planning and execution of live streams and in-person events to promote visibility of women* in science. Providing support on the day of the event, such as administrative logistics, volunteer coordination, and team and public communication.
- Coordination of the Social Media Team. Social media strategy development and implementation. Management of social media platforms (LinkedIn, Twitter, Instagram, Facebook), marketing communication, design and maintenance of the website. Content creation and optimisation, short-form copywriting.

Dragonfly Mental Health

Ambassador and Social Media Coordinator | May 2021 - Present

- Delivering custom talks for undergraduate and postgraduate students as well as members of research and administrative staff. Increasing mental health literacy among academics and providing access to skills trainings.
- Social media management: short-form copywriting and public relations.
- Collaborating with MINDS (www.mindsinacademia.com) to organise MINDS Conference 2021, a virtual conference to provide support and ways for academics to learn about and discuss mental health.

Offspring Magazine the Podcast

Social Media Coordinator | April 2021 - Present

- Coordination of social media and public relations. Establishing social media presence (LinkedIn, Twitter, Instagram), development and implementation of digital strategy. Design and maintenance of the website. Short- and long-form copywriting and content optimisation. Improving podcast accessibility.

EDUCATION

Max-Planck Institute for Biological Intelligence | Munich, Germany

PhD Computational Neurobiology | October 2017 - present

- Max-Planck Phdnet Representative (2019). Participation in regulatory Max-Planck Society (MPS) meetings about MPS-wide policies and Institute Board meetings, facilitating communications between students and directors.
- Working together with the Institute Health Management, promoting intra-institute health initiatives and creating and coordinating community events.
- Co-creator and organiser of Wire Up 2019, an MPS-wide three-day conference to facilitate networking and knowledge exchange among Neuroscience PhD students.
- Member of MPS Mental Health Collective (2021-2022). Promoting mental health via initiatives and MPS-wide events, such as Mental Health Awareness Week.
- Secretary Workgroup member (2020-2022). Supporting Max-Planck institutes in joining Max-Planck Phdnet and carrying out elections.

GSN, Ludwig-Maximillan University | Munich, Germany

MSc Systemic Neuroscience | October 2015 - October 2017

- Course Representative. Organisation and support of Interview and Introduction weeks. Representing students in faculty Board meetings.
- Kids' Brain Day organiser. Designing and implementing a programme to introduce elementary school children to Neuroscience.
- Course tutor and co-founder of the Mentoring Programme for MSc and PhDs.

Newcastle University | Newcastle, UK

BSc Mathematics and Psychology | September 2012 - July 2015

High School No. 45 | Moscow, Russia

International Baccalaureate Diploma | October 2010 - July 2012

5.4 Affidavit

Eidesstattliche Versicherung/Affidavit

Nadezhda Pirogova

(Studierende / Student)

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation Neural Correlates of Contrast Normalization in the Drosophila Visual System selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation Neural Correlates of Contrast Normalization in the Drosophila Visual System is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

München / Munich

21.12.2022

(Datum / Date)

Nadezhda Pirogova

(Unterschrift / Signature)