MICHAEL DREWS

LINEAR AND NON-LINEAR RECEPTIVE FIELDS IN *DROSOPHILA* MOTION VISION

Graduate School of
Systemic Neurosciences
LMU Munich

Dissertation der
Graduate School of Systemic Neurosciences der
Ludwig-Maximilians-Universität
München

3rd June 2019

Michael Drews

Linear and non-linear receptive fields in Drosophila motion vision

This dissertation was submitted to LMU Munich in fulfillment of the requirements for a PhD at the Graduate School for Systemic Neurosciences.

München, 2019

CONTACT
☞ drews@neuro.mpg.de

First reviewer and supervisor
Prof. Dr. Alexander Borst

Second reviewer
Dr. Ruben Portugues

Date of submission
3rd June 2019

Date of defense
23rd September 2019

SUMMARY

The perception of visual motion is a core task to be solved by the visual system of any sighted animal. Visual motion is the displacement of an image or parts of an image that is perceived by an observer. It implies physical movement, whether it be ego-motion, the motion of an object or another behaving organism in vicinity. The specific composition of local motion cues over a scene, so-called optic flow, indicates the direction and speed of ego-motion or rotations around different body-axes. Motion parallax, relative velocity differences between objects, can be used to determine distances. And expanding flow-fields can warn an animal about an imminent collision or the approach of a predator. No matter how complex the behavioral repertoire of a given species may be, the correct interpretation of visual motion is fundamental to the successful execution of locomotion, flight maneuvers, escape responses, mating rituals and many more.

In visual neuroscience, elementary motion detection is a classical problem of neural computation, i.e. the way neural circuits extract meaningful information from sensory input. The displacement of a visual object over time leads to spatio-temporal correlations in the incoming light intensity at the retina, which is the basic sensory gateway for optic signals in most animals. It is not trivial how these spatio-temporal correlations could be exploited by the successive neuronal circuitry in order to generate the percept of motion. Yet, it is a clean-cut problem that can be posed in a quantitative manner. As such it lends itself particularly well not only to mathematical and computational exploration, but also to biological circuit analysis.

The insect visual system has long served as a model system for gaining insight into this question. One reason is that insects show a largely stereotypical behavior to the presentation of visual motion. Flies, for example, always turn with the direction of motion when confronted with a moving image. This so-called optomotor response has been reproducibly shown in various insect species. It is believed to serve as a compensatory reflex to stabilize image velocity on the retina during flight or locomotion in order to counteract external disturbance, such as a wind gust. In addition, early electrophysiology studies discovered the existence of wide-field direction selective neurons in the fly visual system that are the output of the animal's motion vision circuitry. Based on thorough analysis of both behaviour and neurophysiological findings, hypotheses about the nature of the mechanism of elementary motion detection could be made and circuit models could be formulated. With the advent of genetic methodology primarily in the fruit fly *Drosophila melanogaster*, new methods for measuring neuronal activity with sub-cellular resolution as well as circuit manipulation tools have emerged. With these tools at hand it seems finally within reach to test the proposed circuit models directly and to pinpoint the biological counterparts of their algorithmic structure.

During my doctoral studies, I applied some of these tools to create a more complete picture of the individual steps of signal processing occurring between the retina and the first direction selective cells in *Drosophila melanogaster*, so-called T4 and T5 cells. These insights do not only shed light on the cellular implementation of elementary motion detection, they also address questions about which computational strategies the visual system adopts in order to adapt to the requirement of robustness when confronted with complex natural environments far from laboratory conditions. Finally, in the context of *Drosophila* mating behavior, I also investigated neural algorithms relying on an established motion vision system to enable higher visual feature extraction such as figure-ground discrimination.

My findings are presented chronologically in the cumulative thesis at hand. Three of the four manuscripts presented in this thesis were published in peer-reviewed journals and the fourth one is currently subject to peer-review.

In the first study, we systematically mapped all columnar presynaptic partners to T4 and T5 cells in order to describe their spatial and temporal filtering characteristics. Previously, electron-microscopy studies had revealed the presynaptic partner cells of T4 and T5 cells. Other studies had partially described filtering properties of some of those input cells, but using different methods and varying stimuli. We provided a complete and consistent description of all spatial receptive fields and the temporal filter bank formed by those neurons. We found them to be ideally suited to implement temporally asymmetric filtering, a crucial element of computational motion detector models. From that we could deduce a probable anatomical arrangement of those cells along the T4 dendrite, which has now proved correct. By pharmacological intervention, we could show that behavioral state modulation acts already on the inputs of T4, demonstrating impressively the ecological need for adaptive neuronal circuitry already in the early visual system.

In the second study, we investigated how basic visual cues are relayed and further processed for use in complex behaviors. Specifically, we identified a neuronal pathway that enables male fruit flies to closely track and follow a possible mating partner during their courtship ritual. We could show that LC10 neurons relay visual evidence about the presence of small passing objects to the central brain of the fruit fly, possibly towards the courtship circuitry. Intriguingly, their selectivity for small moving targets arises from detecting motion discontinuities between neighboring areas in visual space. This work underlined the fundamental necessity for establishing fully functional motion vision in the early visual system for extraction of behaviorally relevant visual information.

The third study advanced towards a more realistic understanding of the processing cascades in the motion vision circuitry considering also the molecular aspects of neuronal signaling. First, the study confirmed the glutamatergic identity of Mi9 cells, one of the T4 input neurons. Second, while in the first study we only described calcium response properties, here we applied the recently developed glutamate-sensor iGluSnFR to observe directly the transmitter release of glutamatergic neurons. Due to its fast kinetics, we could confirm for Mi9 the temporal filtering properties that we

were only able to assess indirectly by deconvolution from calcium signals in the first study.

Finally, we set out to complete our picture of image processing in fly motion vision regarding non-linear, adaptive phenomena. A drawback of correlation-type motion detectors as encountered in the fly is that they are very contrast-dependent. This makes them unreliable as velocity estimators when confronted with the complexity of natural imagery. Surprisingly, in the last study we were able to demonstrate that this is not the case in the biological system. Using calcium imaging, we probed the circuit for classical mechanisms of gain control. Indeed, presynaptic of T4 and T5, we found evidence for divisive normalization, a widespread phenomenon for gain control, well described in mammals as well as other sensory modalities throughout the brain. Via blocking experiments, we could identify feedback-loops within the circuit as the implementation of this mechanism. Using an unbiased modeling approach by harnessing convolution neural network technology, we could delineate general computational principles substantial for robust motion vision. These considerations nicely paralleled our neurophysiological findings in the fly visual system.

Taken together, the findings presented in this thesis broaden our knowledge of the fly motion vision circuitry and deepen our understanding of the computational solutions nature has equipped the brain with.

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

Among all senses, vision is a particularly complex function that is carried out by our brains. At any moment in time, visual information streams in through our eyes and has to be structured and decoded correctly by the subsequent neuronal circuitry. Given specific patterns of illumination or distribution of light intensities across our retina, our brain manages to decipher relevant information and recognize objects, faces, interpret colors, detect moving objects, identify their direction and many more. We are not aware of this constant process because we are only confronted with the results of these ongoing computations: our perception.

Besides humans, the vast majority of animals relies on vision to guide behavior and thus increase their chances of survival. In this thesis, I study visual processing in the context of motion vision in the fruit fly, *Drosophila melanogaster*, as a model organism. Although natural habitats, anatomy, and evolutionary specification can vary drastically among species, any visual processing system, even an artificial one, faces the same challenges imposed by the structural and statistical properties inherent to natural images (Simoncelli and Olshausen, 2001). From a computational perspective, therefore, algorithmic solutions to general problems of visual information processing often show astonishing analogies (Carandini and Heeger, 2011; Borst and Helmstaedter, 2015).

The fruit fly lends itself particularly well to the investigation of motion processing circuits. First, this is due to its stereotypical behavior which allows for a reliable read-out of motor reflexes that are guided by visual motion. Second, unraveling the architecture and understanding the function of a neural circuit seems more feasible in a fly brain which contains orders of magnitude fewer neurons than most vertebrate model organisms. Finally, the availability of an immense range of cell-targeted genetic manipulation tools combined with neurophysiology, imaging techniques and behavioral analysis offers a compelling toolbox and unparalleled opportunities for circuit analysis on a cellular level.

In the following, I will provide the background for the articles presented in the main part of this thesis. I will begin by covering theoretical aspects of visual and motion processing in general, provide a summary of state-of-the art research in the fly visual system and finally introduce the experimental techniques which have been used in this work.

1.1 VISUAL PROCESSING

What is vision? Biological visual systems have evolved to allow an organism to detect light signals and to extract information about the environment. The fact that optical signals are made of light which is emitted or reflected by

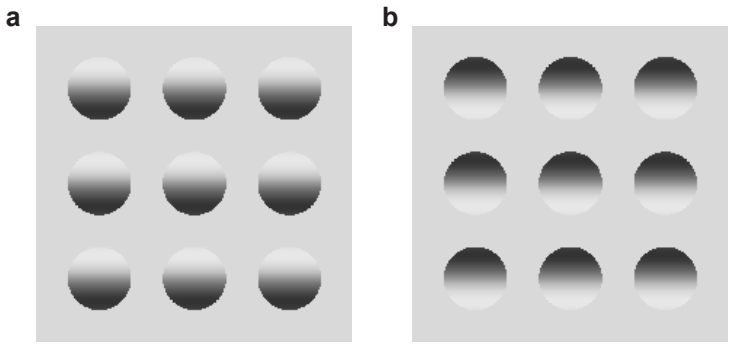


Figure 1: Shading Illusions | **a**, The objects appear as convex because the brain assumes that light usually comes from above. **b**, If the shading is reversed, objects appear as concave. Graphics adapted from Ramachandran (1988).

surrounding objects means that they contain information about the position, size, type and many other features of these objects. In other words, structure and regularity in the physical world entails structure and regularity in optical images of that world. Visual systems have been designed to process these images and to recover that structure from them.

In vision research, the visual system is viewed not as a general image processing machine but as a system that is very well adapted to the particular problem of processing sequences of natural images. In the course of evolution, the brain has acquired an enormous amount of prior information about what kind of data it can expect to receive from its sensory input organs. An instructive example for this are shading illusions as depicted in Figure 1. While objects appear as convex circular bumps when they are shaded brighter at the top than at the bottom, they seem to be concave when the shading is reversed. Our brain makes an inference about the three-dimensionality of these objects based on the direction of this shading gradient and under the assumption that, if there is no other obvious light source present, light usually comes from above.

The idea that our visual system uses prior information about the natural world has been first put forward by von Helmholtz (1867) under the term "unconscious inference" in the context of such optical illusions. We can find that idea again in generative models of perception (Dayan et al., 1995) which suggest that what the visual system does is trying to match its observations with an internal model of the world. More specifically, it can also be applied to build a theory for understanding early visual processing, as for example in the mammalian retina or in the *Drosophila* visual system. Normative theories like redundancy reduction or efficient coding (Attneave, 1954; Barlow, 1961) try to transcend descriptive models of neuronal processing. They are able to link the functional properties of early visual neurons to general coding principles and the requirement for an optimized neural representation of naturalistic stimuli in the brain.

In the following, I will focus on some aspects of these theories and illustrate how general statistical properties of natural input stimuli are reflected in mechanisms of neural processing.

1.1.1 Natural images

Image formation in any visual system, whether a biological or an artificial one, is based on light projection. The result is a two-dimensional image of the physical world. Obviously, the images produced by observing the natural world are countless and as diverse as nature itself. It may therefore come as a surprise that the set of all possible natural images is actually by far smaller than the set of possible images in general. Just consider a grayscale image patch of 100x100 pixels with each 256 possible intensity values. The set of all possible images then comprises $256^{10,000}$ images since each pixel's intensity value can be chosen independently from its neighbor. This is an immense number. However, many of those randomly generated images will appear simply as fuzzy and noisy image patches without any meaningful structure to our eyes. In contrast to that, natural images possess many obvious structural elements such as edges, gradients, textured areas and so on. Despite their diversity, natural images therefore share certain fundamental features that are statistical characteristics of the class of natural images. A description of a few of the main statistical properties of natural scenes is given in the following.

Luminance and contrast

The first order statistic of natural images is the distribution of individual pixel intensity values across the set of natural scenes. Early work has consistently confirmed that this local luminance distribution is positively skewed (Figure 2a,b) (van der Schaaf and van Hateren, 1996; Brady and Field, 2000; Geisler, 2008). This means that there are more pixels which are darker than the average luminance, but there are a few pixels that are a lot brighter than one would expect from a Gaussian distribution with the same standard deviation. This distribution becomes symmetric after a logarithmic transform of the image.

The contrast of an image is not rigorously defined. For artificial stimuli like gratings or bars, one might use the Michelson or the Weber contrast which are ratio-metric measures derived from minimum and maximum intensity values. For natural scenes, it makes more sense to look at the root mean square contrast, i.e. the standard deviation, of pixel intensities in a local patch (Figure 2c). Contrast is then a measure of local intensity variance which is also positively skewed. Contrast and luminance are largely statistically independent, although they correlate over small distances and in sky regions where luminance is generally high and contrast is generally low (Frazor and Geisler, 2006; Geisler, 2008).

Spatial structure and power spectrum

The second order statistics of natural images take into account pairwise correlations between pixels. Such correlations can be characterized using the autocorrelation function or the power spectrum. Power spectra of natural images typically follow an exponential power law $1/f^n$ with an exponent $n \approx 2.0$ (Figure 2d) (Field, 1987; Ruderman, 1994). This means that low spatial frequency components, i.e. long spatial wavelengths corresponding

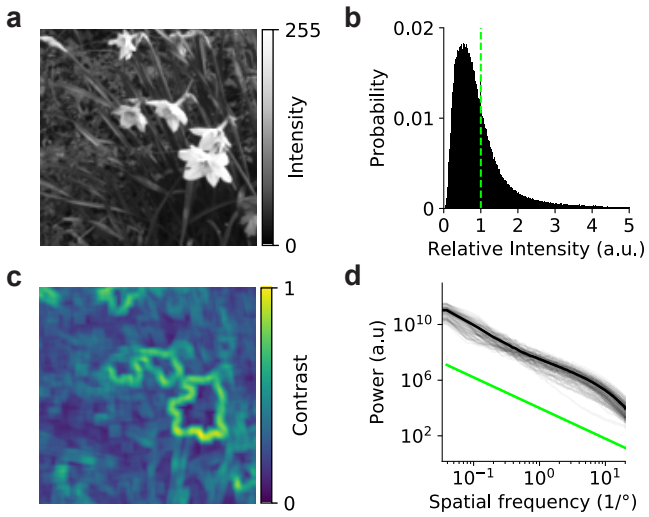


Figure 2: Natural image statistics | **a**, Representative natural image patch. **b**, Histogram of pixel intensities over 1000 natural images. The intensity distribution of natural images is strongly skewed towards brighter than average intensity (green dashed line). **c**, Local contrast map for the image patch in a, estimated by taking the root-mean-square contrast in a 50x50 pixel sliding window. **d**, Power spectrum of 100 natural images (gray lines, black line indicates the mean) on double-logarithmic axes. Green line shows a power law with exponent -2.2 for reference. Images taken from van Hateren and van der Schaaf (1998).

to luminance changes over large distances, dominate natural images. Scaling the argument of a power law function gives a power law function with the same exponent. Hence, this is another way of saying that natural scenes are scale invariant. Scale invariance means that zooming in or zooming out of an image does not change the statistical properties of that image. Efforts trying to provide an explanation for these self-similar properties have argued that natural images are in general composed of many statistically independent objects of different sizes which are placed at different distances and this does not depend on the position of the observer (Ruderman, 1997).

Another consequence of the exponential power-law is that the autocorrelation function of natural scenes is also an exponential. Thus, neighbouring pixels in natural images tend to have more similar values the closer they are to each other. This intuitively makes sense because the probability is high that neighbouring pixels belong to the same object or structure.

A closer analysis of the two-dimensional power spectrum of natural scenes also reveals some statistical anisotropies. In general, there are two peaks in the power spectrum for image components oriented along the horizontal or vertical axis of the image. This is because the horizon or vertical structures like tree branches or grass tend to be dominant in most natural image datasets (van der Schaaf and van Hateren, 1996; Dyakova and Nordström, 2017). For the purpose of this thesis, however, such detailed description of natural image statistics will not be relevant.

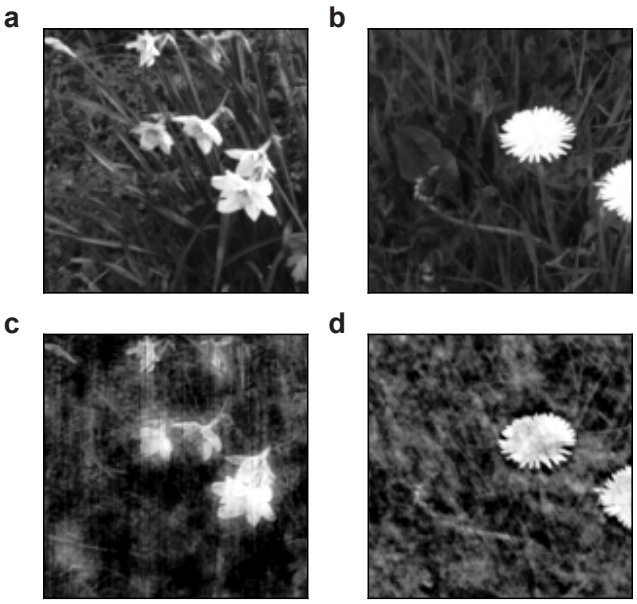


Figure 3: Phase hybrid images | **a,b**, Original natural image patches. **c,d**, Corresponding hybrid images with the same phase spectra, but amplitude spectra swapped. Images taken from van Hateren and van der Schaaf (1998).

Higher-order statistics

The statistical properties of natural images are far more complex than what can be captured by examination of first and second order features. Above, we have considered only the intensity distribution and the power spectrum. A large part of the relevant image content of natural scenes, however, is contained in the phase spectrum. This becomes clear from looking at hybrid images where the amplitude spectra of two images have been swapped in Fourier space, but the phase spectra are not manipulated (Figure 3). Although hybrid images seem heavily distorted, we are still able to identify the same object boundaries and the image content appears to be unchanged to our eyes (Tadmor and Tolhurst, 1993). This must be due to the information contained in the phase spectra.

Observations like these can by explained by higher-order statistical features, an example of which is co-occurrence of edges of similar orientations. It has been shown that there is an increased probability of finding similarly oriented edges in the vicinity of a given edge because contours in natural images tend to be elongated and parallel structures tend to be located close to each other (Geisler et al., 2001).

Even uncorrelated image features can still be statistically dependent on each other. Schwartz and Simoncelli (2001) showed that the response strength of a local edge filter is correlated to the variance of a differently oriented edge filter with uncorrelated output. This is related to the finding that also local contrast, i.e. local variance, seems to be scale invariant and thus spatially correlated (Ruderman, 1994). More intuitively, just as the spatial extent of objects and other structures in natural images leads to correlations between neighbouring pixels, also image textures with similar contrast seem to occupy extended regions in visual space which results in

correlations of local contrast. Instead of intensity, the statistical dependence here affects the variance of pixel intensity in a given area. This kind of higher-order statistics regarding local contrast has strong implications for cortical and retinal processing, as we will see.

1.1.2 Redundancy reduction

As pointed out above, natural scenes exhibit strong spatial correlations (see Section 1.1.1). Therefore, knowledge about the intensity of an arbitrary pixel in a natural image implies some knowledge about the intensity of close-by pixels: although this does not have to be the case, chances are high that they have very similar intensities. In information theoretic terms this means that a pixel-based representation of natural images is a *redundant* representation (Shannon, 1949). It follows that one can store all image information without explicitly coding for the absolute intensity of each pixel because it is possible to infer some information from neighboring pixels. In theory, we could find a more efficient representation and compress the image by taking advantage of what we know about pairwise pixel correlations.

The notion that the visual system aims to reduce the redundancy of natural images in order to find a more efficient neural representation was originally proposed by Attneave (1954). Similarly, the *efficient coding hypothesis* by Barlow (1961) states that it is a central goal of all sensory coding to minimize the amount of nervous activity required to represent sensory information.

Without further formalizing these concepts in mathematical terms, in the following I will outline how the autocorrelation of natural images is exploited by early visual processing mechanisms in order to reduce the redundancy of natural images. The efficient coding hypothesis leads to the prediction of optimal local processing filters for natural scenes. These predicted filters strongly resemble linear receptive fields of early visual interneurons found in the mammalian retina, but also in the fly lamina. As we will see in Manuscript 1 in the main part of this thesis, this is also reflected in the spatial filtering characteristics of neurons in the fly medulla. In the following, I will introduce the concept of a *receptive field* and delineate how the spatial structure of receptive fields in early vision can be linked to the principle of redundancy reduction.

Receptive fields

The *spatial receptive field* of a visual neuron is quite generally defined as the area in visual space, in which the activity of that neuron can be modified by visual stimuli. This definition can be extended in the temporal domain to the time window in the past when presentation of visual stimuli within the spatial receptive field had an impact on the instantaneous activity of the neuron.

Classically, the spatial receptive field of a neuron was investigated by means of flashing light stimuli at different positions, or by other types of stimulus protocols correlating local luminance changes with neuronal activity. While originally the concept of a receptive field was defined by

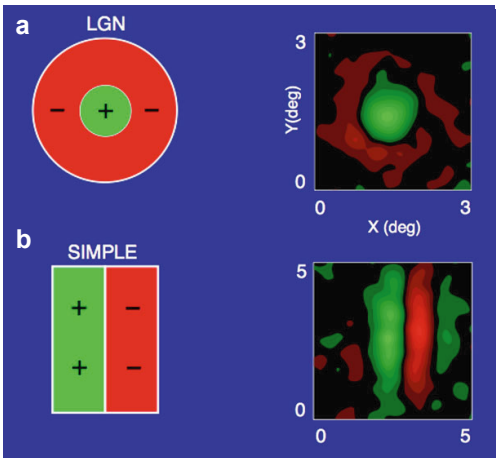


Figure 4: Spatial receptive fields | a, Schematics of a center-surround antagonistic receptive field (left) and corresponding receptive field measurement from an LGN neuron (right). **b**, Schematics of orientation-selective edge filter (left) and corresponding receptive field measurement from a V1 simple cell (right). Graphics adapted from Lindeberg (2013) / CC BY.

observing only positive correlations between illumination of a certain area in visual space and neuronal discharges (Hartline, 1940), it was soon acknowledged that receptive fields can have excitatory and inhibitory subregions (Kuffler, 1953; Hubel and Wiesel, 1959). We call them *ON*-subregions, when raising light intensity in that area increases neuronal activity and reducing light intensity decreases activity, and *vice versa* for *OFF*-subregions. The summation field that results from the specific spatial composition of such ON- and OFF-subregions is called the *linear* receptive field of a neuron (Figure 4).

Spatial receptive fields can take on arbitrarily complex shapes. Often, specific receptive field characteristics can be attributed to a class of neurons in the brain. Since in visual systems most neuronal structures represent visual space in a retinotopic way, the underlying computational operation that is implemented by a class of neurons with a characteristical receptive field can be interpreted as a convolution of the stimulus with a spatial filter. The filter kernel is given by the linear receptive field. For example, neurons in the *laterate geniculate nucleus (LGN)* in cats and primates possess approximately symmetrical center-surround receptive field configurations which are characterized by a center region (ON or OFF) that is suppressed by an antagonistic surround region of the opposite polarity (Figure 4a) (DeAngelis et al., 1995; Bonin et al., 2005). The image processing operation implemented by such receptive field configurations corresponds to high-pass filtering or an approximation of the spatial derivative of an image. Neurons with such receptive fields respond strongly if the difference in stimulus intensity between center and surround regions is large. In other words, they carry information about local contrast rather than absolute local light intensity. A different class of neurons, *simple cells* in the cortical area *V1*, possess elongated spatial receptive fields with parallel ON- and OFF-subregions (Figure 4b) (DeAngelis et al., 1995; Lindeberg, 2013). Due to their

specific spatial structure they act as local edge filters that respond selectively to correspondingly oriented image contours.

However, neurons can implement much more complex computations than simple linear image filtering. For example, they can be subject to non-linear modulation by higher-order stimulus features (Carandini and Heeger, 2011) or they can perform inherently non-linear operations such as motion detection (Borst and Helmstaedter, 2015). In such cases, it is not as straightforward to depict the spatial receptive field as a summation field of spatially separable ON- and OFF-subregions. We then must return to the original definition which defines the spatial receptive field merely as the area in which suitable stimuli can modify the activity of a neuron.

Decorrelation and whitening

Having established the term of a receptive field, we now return to the idea of efficient coding of natural images. I have mentioned above that, in information theoretic terms, spatial correlations between nearby pixels in natural scenes represent a source of redundancy. A more efficient representation of natural images would contain all image information but without coding for this redundant part. One way to reduce redundancy would therefore be to find a way to *decorrelate* the data, i.e. to reduce the spatial correlations of natural scenes.

The *whitening theory* proposes that early visual neurons perform this decorrelation by spatial filtering. The power spectrum is related to the autocorrelation function via the Fourier transform. Since the autocorrelation of an uncorrelated signal is a delta function, which has a flat power spectrum in Fourier space, decorrelating of a signal means flattening its power spectrum, one says also *whitening*. Mathematically, whitening of a dataset can be performed using *principal component analysis (PCA)*.

Principal-component analysis

PCA is a statistical method that is used to find a transformation for a multi-dimensional dataset of possibly correlated observations that changes it into a *linearly* uncorrelated dataset. This method relies on finding the so-called *principal components* which are a set of uncorrelated features that are aligned with the directions of highest variance in the dataset. Because these principal components explain most of the variance in the data, they are also the most informative directions in data space. Expressing the dataset as a linear superposition of these by definition uncorrelated features is therefore a more efficient representation. More mathematical details of PCA are presented in Pearsons' original article (Pearson, 1901) or in standard textbooks about statistics or data science as well as about visual processing (Hyvärinen et al., 2009).

Zero-phase whitening filters

How can we apply PCA to natural images? This is straightforward by expressing an image as an n-dimensional vector with n equal to the number of pixels per image. Applying the PCA algorithm to a dataset of such vectors

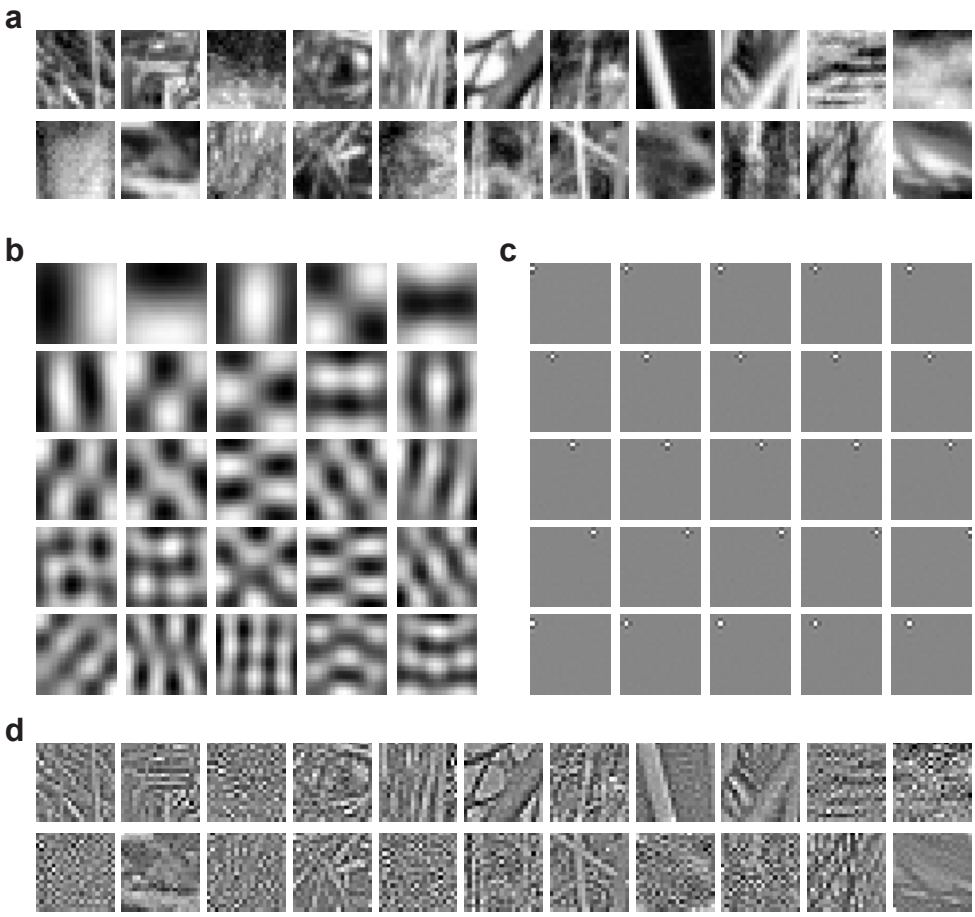


Figure 5: Decorrelation of natural scenes | a, Example 20x20 pixel image patches from natural scenes. **b**, The first 20 principal components of a dataset of such natural image patches. Ordered by variance explained, decreasing from left to right and top to bottom. **c**, 20 example ZCA filters in retinotopical order. **d**, Applying ZCA whitening to the original image patches yields whitened images with zero phase-shift. They appear like differentiated versions of the original images, where large deviations from medium gray occur only at positions of large intensity changes, i.e. high contrast, in the original picture. PCA and ZCA were trained on 20,000 logarithmically compressed randomly chosen image patches from the image database by van Hateren and van der Schaaf (1998).

yields the principal components, ordered by the amount of variance they are able to explain. When performed on a large enough database of natural images, the first principal components look like whole-field sinusoids with larger wavelengths first and then increasingly fine-grained as the variance explained by the component keeps decreasing (Figure 5a,b). This makes sense, since the $1/f^2$ power spectral density tells us that small image frequencies, i.e. large spatial wavelengths, dominate natural images. Therefore, PCA performs a kind of frequency analysis of the dataset (Hancock et al., 1992). Building an image representation based on linear superposition of these principal components, is not very sensitive to the information contained in the phase spectrum because of the large spatial extension of these features. As we have seen however, the phase spectrum carries a lot of perceptually

relevant information, e.g. about the location of object boundaries and edges (Figure 3).

PCA is not the only whitening transformation that exists. In fact, it can be shown that PCA is just one of an infinite number of whitening transformations that can be constructed. We can constrain this number by demanding that the transformation be symmetrical, i.e. that the resulting features are local and produce no phase offset when used as spatial filters. This solution is also called *zero-phase component analysis (ZCA)* (Bell and Sejnowski, 1997). The resulting components are spatially symmetrical filters that have a center-surround antagonistic structure (Figure 5c). They are spatial high-pass filters that attenuate low frequencies and let higher frequencies pass, thereby flattening the power spectrum of natural scenes. Applying these filters to natural images produces essentially differentiated versions of the originals where large positive or negative deviations from zero indicate large differences between neighboring pixel intensities in the original image (Figure 5d). Thus, this is a representation of local contrast instead of absolute pixel intensity. Here, the spatial correlations between neighboring pixels in the original image have been filtered out.

Physiological evidence for whitening

Spatial center-surround filters like those constructed above from a theoretical point of view have often been described in the early visual pathway of different model organisms. Early studies have discovered center-surround receptive fields in the mammalian as well as non-mammalian retina (Kuffler, 1953; Barlow, 1953; Rodieck, 1965). Neurons in the lateral geniculate nucleus (LGN), one level higher up in the mammalian visual processing hierarchy, have been found to possess such receptive field structure as well (DeAngelis et al., 1995; Bonin et al., 2005). Center-surround filtering has been described also in *large monopolar cells* in the early visual system of evolutionary distant flies (Dubs, 1982; Srinivasan et al., 1982). Similar observations have been made in crickets (Honegger, 1980). Overall, these findings suggest that symmetrical antagonistic filtering is indeed a wide-spread functional motif in visual processing across phyla. In Manuscript 1, we will extensively map receptive fields in the fly medulla yielding similar results.

Another type of whitening transform, *independent component analysis (ICA)*, requires the outputs not only to be decorrelated but also to be statistically independent. Applying ICA to natural images yields filters which are semi-local edge detectors and resemble spatial receptive fields of simple cells found in visual cortex (Bell and Sejnowski, 1997).

Do neurons that have been attributed with center-surround spatial receptive fields indeed perform decorrelation of real-world input stimuli? The abundant evidence for such receptive fields in early vision is not a direct proof that this mechanism actually increases the coding efficiency of neurons when confronted with naturalistic input signals. This is a difficult question, since we do not know for which task specifically these neurons have been optimized. Viewing static natural images on a screen is a simplification of the complex nature of real-world sensory input that is generated during animal behavior in a dynamic and constantly changing

environment. One study tried to address this issue by showing natural movies to cats while performing electrophysiological recordings in LGN (Dan et al., 1996). Indeed, they found power spectra of LGN responses to be essentially white, thus confirming the whitening theory. In flies, some studies have predicted the spatio-temporal receptive field structure of large monopolar cells from the efficient coding point of view (Srinivasan et al., 1982; van Hateren, 1992a,b). In fly vision, studies are often put in the context of motion vision which places major relevance also on the temporal filtering characteristics of those neurons, such as biphasic responses with temporal inhibition. In salamander retina, another study confirmed that naturalistic stimuli decorrelate responses of retinal ganglion cells, though they also found that most of this effect can actually not be attributed to linear filtering properties, but to additional non-linear mechanisms (Pitkow and Meister, 2012). However, efficient coding is a general hypothesis which is not restricted to exclusively linear computations. The next section will introduce an important non-linear phenomenon in visual processing.

1.1.3 Contrast normalization

Above, we have seen how the $1/f^2$ power spectrum of natural scenes (see Section 1.1.1) and the hypothesis that visual processing has been adapted to such input statistics in order to optimize its coding strategy, have led to the prediction of center-surround antagonistic receptive fields (see Section 1.1.2). But natural scenes show much more statistical structure than is captured by the second-order statistics, i.e. the power spectrum. One particular aspect is that higher-order features, such as pixel variance or contrast, are subject to statistical fluctuations as well (Ruderman, 1994; Schwartz and Simoncelli, 2001). Contrast fluctuations in natural images also show spatial correlations. For example, sky areas are low-contrast regions which often cover large parts of an image. Other image regions might contain textures which are rich in contrast, for example tree branches against the sky, or also medium contrast regimes, like grassy areas. Additionally, different illumination conditions can have a drastic impact on the contrast of natural visual environments. For instance, objects' boundaries are usually much more pronounced under direct daylight illumination than under dim light conditions.

Divisive normalization

Neurons have a limited dynamic range that is set by biophysical properties. The magnitude of contrast fluctuations in natural scenes, however, can be large and strongly depends on the position in an image and on current lighting conditions. Hence, contrast gain adaptation is a widespread phenomenon in the visual system. Gain control mechanisms dynamically set the input range of a neuron to the range of expected input signals. One way to obtain such estimation for the current input range is to assess the stimulus variance based on its history. This is exploited by temporal gain control mechanisms, commonly found in early visual processing in the retina (Baccus and Meister, 2002; Demb, 2008) and also in the fly (van Hateren, 1997; Harris et al., 2000). In order to avoid long integration times,

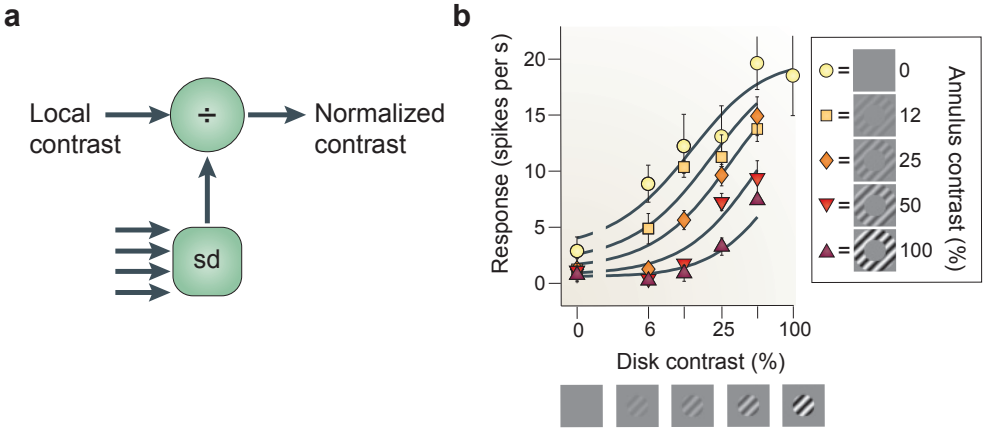


Figure 6: Contrast normalization in primary visual cortex | **a**, Contrast normalization transforms local contrast into normalized contrast by means of division by an estimate of surround contrast. Surround contrast is estimated by a computation approximating the standard deviation (sd) of the stimulus in a surrounding area of visual space. **b**, Responses of a visual cortex neuron to gratings within a central disk surrounded by an annulus of different contrast. Symbols indicate data, black lines are fits of a divisive normalization model. High surround contrast suppresses responses to the center disk. This leads to a shift of local contrast tuning curves towards higher contrasts, effectively modulating contrast sensitivity for the driving stimulus in the center. Graphics taken from Carandini and Heeger (2011).

however, the visual system also employs strategies to estimate local contrast in the spatial domain. In *divisive contrast normalization*, the response of a neuron is effectively divided by an estimate of surround contrast which is formed by a pooling mechanism that spatially integrates rectified or squared local contrast-sensitive units (Figure 6a). The suppressive signal generated by such spatial integration mechanism is closely related to the root-mean-square contrast of an image patch. This kind of gain control has been abundantly described in vertebrate visual processing (Carandini and Heeger, 2011). On an algorithmic level, contrast normalization can be described by the following equation:

$$R = \frac{\sum_i w_i C_i}{k + \sqrt{\sum_k \alpha_k {C_k}^2}}$$

Here, the weights w_i define the linear spatial receptive field of a neuron and the C_i are the input signals, which is local contrast here. After image pre-processing (e.g. through whitening filters), input signals correspond to local contrast rather than local luminance. This linear response is then divided by an estimate of root-mean-square contrast in the surround computed over a separate spatial summation field defined by the weights α_k. The constant k prevents division by zero.

This equation predicts contrast sensitivity to depend on the surround contrast of a locally contrast-sensitive neuron (Figure 6b). Baseline sensitivity is set by the constant k which determines the half-saturation threshold of the neuron. As the contrast in the surround is increased, the suppressive field α_k is stimulated which reduces the gain for the driving stimulus in the center of the linear receptive field. On a logarithmic axis, this is as if contrast tunings

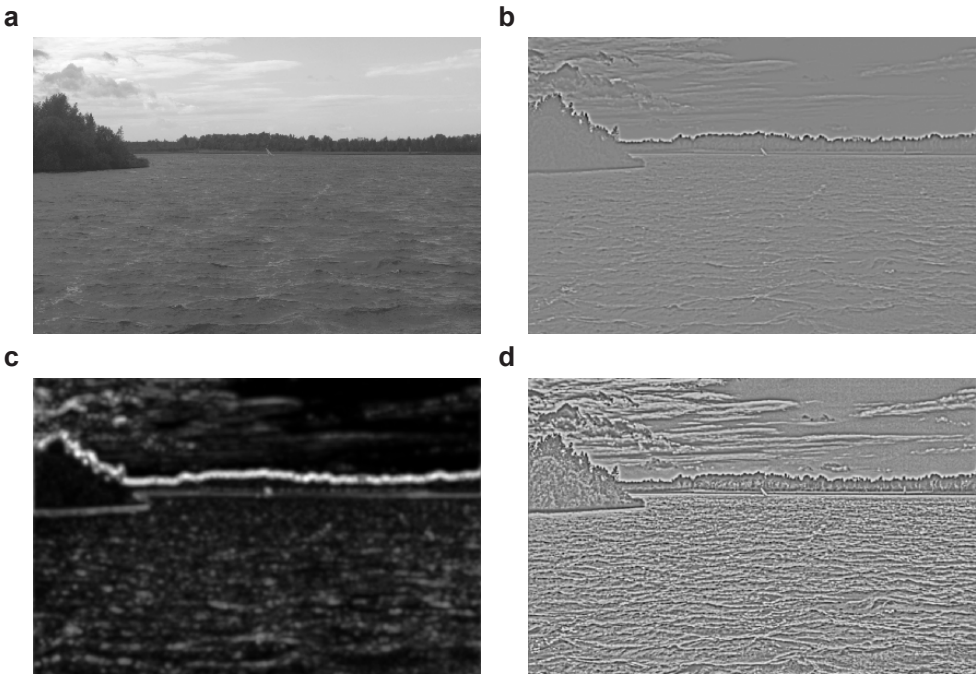


Figure 7: Contrast normalization as image processing operation | **a**, Original gamma-corrected picture, taken from van Hateren and van der Schaaf (1998). **b**, Image after linear center-surround filtering by a difference of Gaussians filter. Bright and dark pixels now indicate strong local ON- or OFF-contrast, respectively. Large areas of the image have relatively low contrast while clear contours are mostly located along the horizon. **c**, Average local contrast estimated by the root-mean-square contrast in a Gaussian window of the same width as the surround of the linear filter. **d**, Image after local contrast normalization by division of the local contrast representation in b through average local contrast as calculated in c plus a small constant. Edges and contours are now clearly visible across the whole image. This representation does not indicate local contrast anymore, but a form of normalized contrast.

are shifted horizontally towards higher contrasts. The contrast sensitivity of the neuron is thus dynamically adjusted to encode the range of contrast values that is expected from the estimated average contrast in the vicinity of the receptive field center.

Similarly to how linear center-surround filtering transforms local luminance into local contrast, contrast normalization transforms local contrast into a relative measure of local contrast, normalized to the average contrast in the surround. From an information theoretic point of view this reduces contrast correlations in natural scenes, further improving the coding efficiency of the neural representation of natural scenes (Brady and Field, 2000; Schwartz and Simoncelli, 2001; Wainwright et al., 2002). More intuitively, the operation to some degree equalizes the contrast variability across an image (Figure 7). Contrast normalization is suppressive in high contrast regions and amplifies signals in low contrast regions which overall leads to a more uniform representation of local contrast. This enhances the detectability of object boundaries and contours across the image regardless of the local contrast conditions in each subregion.

Apart from that, normalization as a neuronal computation has been suggested to fulfill multiple functions in sensory processing, such as adjustment of neural sensitivities, but also inducing invariance of neural representations with respect to uninformative stimulus dimensions, improving stimulus discriminability and more (Carandini and Heeger, 2011).

Physiological evidence for contrast normalization

Contrast normalization and related phenomena have been investigated extensively in the cortex (Heeger, 1992; DeAngelis et al., 1992; Carandini et al., 1997; Carandini and Heeger, 2011). In addition, it has been shown to act throughout the mammalian visual processing hierarchy, in the LGN (Freeman et al., 2002; Solomon et al., 2004; Bonin et al., 2005) as well as in the retina (Shapley and Victor, 1978; Demb, 2008). Although we introduced normalization here in the context of visual processing, the same computation has been described in other sensory modalities. These include for example auditory processing in the ferret (Rabinowitz et al., 2011) and the fly olfactory system (Olsen et al., 2010), where it has been shown to be beneficial in terms of coding efficiency (Luo et al., 2010). Normalization has therefore been suggested to serve as a so-called *canonical computation* (Carandini and Heeger, 2011) that is repeatedly applied across different brain regions, modalities and species to solve similar problems in a variety of contexts. In this thesis, I will complement this body of work and present evidence for contrast normalization in the *Drosophila* visual system where it is critical for robust motion vision in natural environments (see Manuscript 4).

1.1.4 Motion detection

Above, we have seen how the input statistics of natural stimuli are exploited by computational mechanisms of the visual system (see Section 1.1.2 and Section 1.1.3). However, we only looked at static natural input stimuli without really considering the fact that real-world visual input is, of course, subject to constant change. Most of the changes that take place in our environment are due to motion of other objects or self-induced ego-motion. In physics, motion is defined as the displacement of a body in a given time interval. Consequently, motion of an object is reflected in a temporal displacement of its visual appearance on our retina. Object detection takes place at higher-level stages of visual processing, while at the level of the retina visual objects are not yet clearly defined entities in the neural representation. Since detecting the direction of motion is, however, such a critical task for an animal to react quickly and without delay to changes in its immediate environment, visual systems have come up with solutions for motion detection based on retinal input directly. The goal of this section is to introduce some basic computational principles that guide elementary motion detection in neural networks.

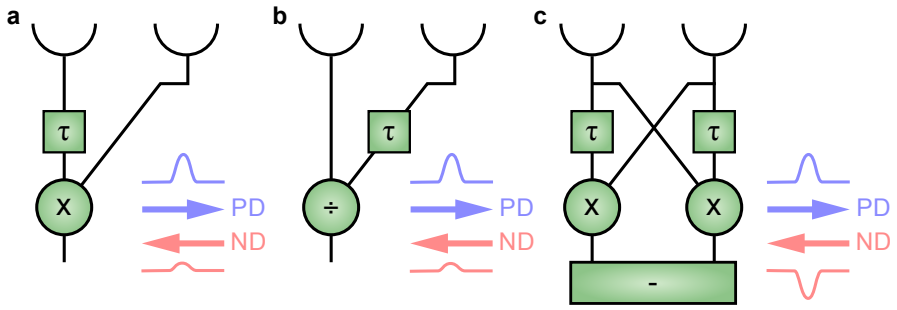


Figure 8: Two-arm models of motion detection | **a**, Hassenstein-Reichardt-type half-detector. Signals are non-linearly enhanced (here by multiplication), when a stimulus moves in the preferred direction (PD). Stimuli traveling in the null direction (ND) elicit no or only small responses. **b**, Barlow-Levick-type half-detector. Signals are non-linearly suppressed (here by division), when a stimulus moves in the ND. Stimuli moving in PD are not suppressed, so PD responses are larger than ND responses. **c**, Full Hassenstein-Reichardt-detector. Subtraction of two mirror symmetric half-detectors gives positive responses to PD motion and negative responses to ND motion.

Hassenstein-Reichardt detectors

From a theoretical point of view, visual motion is reflected in spatio-temporal correlations of the visual input. The image of a moving object which is detected at a given position on the retina will elicit responses in spatially offset photoreceptors some time interval later. A neuron is called direction-selective when it can detect the direction of these spatio-temporal correlations, i.e. when it responds to motion in one direction but not (or less) to motion in the other direction. This defines some minimum requirements that are necessary for elementary motion detection (Borst and Egelhaaf, 1989):

1. **Spatial offset**
   At least two spatially offset inputs are required to detect motion. Observing a single point in space is not informative about motion in any direction.

2. **Temporal asymmetry**
   The input channels have to be asymmetric in their temporal filtering properties. If the detector was symmetric, its mirror-image would produce the same output signal. Hence, the detector could not respond in a direction selective way.

3. **Non-linear interaction**
   The input channels need to interact in a non-linear way. Otherwise, the time averaged output signal would be the same for opposite directions of motion.

If we consider the case of motion detection from two spatially offset input channels, this leads to the distinction of two alternative mechanisms of motion detection: First, based on behavioral studies on the snout weevil *Chlorophanus viridis*, a mechanism involving multiplicative interaction

between the two input channels was suggested, the so-called *Hassenstein-Reichardt (HR)* detector (Hassenstein, 1951; Hassenstein and Reichardt, 1956). Here, the channel that is stimulated first when a motion stimulus passes in the detector's preferred direction (PD) has a longer transmission time constant in comparison to the other channel (Figure 8a). This compensates for the time that the stimulus takes to bridge the spatial distance to the second channel and thus leads to a temporal overlap of both signals at the multiplication stage. Here, the two signals interact non-linearly which leads to an overall strong output response. If the stimulus travels in the opposite, the null direction (ND), the signals have less of a temporal overlap and so less interaction energy is released. The direction selectivity of this type of detector is therefore based on an *enhancing* non-linearity which amplifies responses to PD stimuli. This model is called a full HR detector if two such mirror-symmetric subunits are subtracted from each other (Figure 8c). Adding the subtraction stage leads to direction-opponent responses: full HR detectors produce positive output signals for PD stimuli and negative output for ND stimuli. Such fully symmetric HR models have successfully been applied to explain behavioral responses of flies and other invertebrates (Hassenstein and Reichardt, 1956; von Fermi and Richardt, 1963; Götz, 1964) as well as to account for electrophysiological findings in motion-selective neurons in the fly (Haag et al., 2004; Joesch et al., 2008; Borst et al., 2010).

Barlow-Levick detectors

While non-linear interaction is a prerequisite for motion detection, the exact nature of the non-linearity can be very different from the HR detector. An alternative model, the so-called *Barlow-Levick (BL)* detector replaces the multiplicative stage of the HR detector with a *suppressive* non-linearity. While originally this suppression was modelled as asymmetric inhibition in order to account for direction-selectivity of retinal ganglion cells in the rabbit retina (Barlow and Levick, 1965), here I depict it as a divisive interaction in order to accentuate the discrepancy between the two models (Figure 8b). In the BL detector, the temporally delayed arm is the second arm of model. That way, temporal overlap is maximized for stimuli moving in the ND. As a consequence, a strong suppressive interaction takes place and renders responses to ND stimuli small. On the contrary, if the stimulus travels in PD, the signal that is elicited in the first receptor can pass without being suppressed by the second channel because the signal from the second arm arrives "too late". Hence, direction selectivity in the BL detector relies on suppression of ND stimuli. These two opposite mechanism of motion detection are often referred to as *PD enhancement* and *ND suppression*.

Hybrid detectors

Although historically, the HR detector and the BL detector have been treated as mutually exclusive alternatives, recent findings in *Drosophila* suggest that both mechanisms could actually be implemented simultaneously on the dendrites of direction-selective T4 and T5 cells (Haag et al., 2016). Conceptually, this can be represented by fusing the two detector subunits, one implementing PD enhancement and one based on ND suppression,

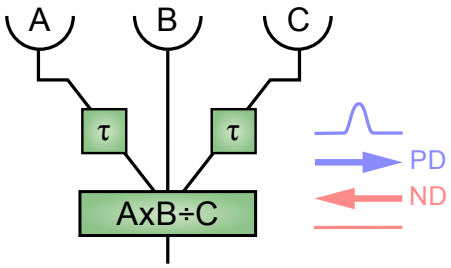


Figure 9: Three-arm hybrid detector for motion detection | Combining HR-type and BL-type motion half-detectors into one detector with three input channels yields an elementary motion detector with increased direction selectivity.

with the same preferred direction into one motion detector that gets input from three channels (Figure 9). This three-arm hybrid detector combines the two mechanisms to generate a more direction-selective signal. It has been suggested to account for the high degree of direction selectivity observed in T4 and T5 cells in the fly (Haag et al., 2016). This kind of redundant implementation of complementary algorithmic solutions is often observed in neurobiology. In fact, there is also evidence that both, asymmetric excitation as well as asymmetric inhibition, play a role in motion processing circuits in the vertebrate retina (Briggman et al., 2011; Baden et al., 2013; Ding et al., 2016; Vlasits et al., 2016; Mauss et al., 2017).

Other motion detectors

The models of motion detection discussed above are also called *correlation-type motion detectors* because they are based on the extraction of spatio-temporal correlations from an array of visual input signals. Responses of such correlation-type models depend not only on temporal but also on spatial properties of the input stimulus. When stimulated with sinusoids, they are tuned to the contrast frequency of the stimulus rather than its absolute velocity (Egelhaaf et al., 1989; Srinivasan et al., 1999). However, when stimulated with a mix of spatial scales such as pervasive in natural scenes, they show velocity dependence (Dror et al., 2001).

Another computational approach, the *gradient detector*, which was originally proposed in the context of machine vision (Limb and Murphy, 1975; Fennema and Thompson, 1979), tries to overcome this pattern dependence. In the gradient detector model, local stimulus velocity is estimated by calculating not only the temporal gradient $\frac{dI}{dt}$ of the input signal but also the spatial gradient $\frac{dI}{dx}$. Since the temporal gradient can also be expressed as the product of velocity v and the spatial gradient,

$$\frac{dI}{dt} = v \cdot \frac{dI}{dx}$$

it follows that

$$v = \frac{dI}{dt} \div \frac{dI}{dx} = \frac{dx}{dt}$$

Hence, stimulus velocity is given as the ratio between the two gradients, independently of the spatial structure of the stimulus. Although this might

seem beneficial in many contexts, is has been argued that gradient detectors might be disadvantaged in comparison to correlation-type detectors when confronted with noisy signaling conditions (Potters and Bialek, 1994; Borst, 2007).

Another type of motion detection model that is frequently employed in vertebrate vision and human psychophysics is the so-called *motion energy model* (van Santen and Sperling, 1984; Adelson and Bergen, 1985). In models of this type, direction selectivity is generated by convolution of a space-time representation of the stimulus with a spatio-temporally tilted linear filter and subsequent thresholding, squaring or another type of non-linearity. However, it can be proven that the output of the motion energy model is equivalent to the output of the HR detector, despite their different internal structures (van Santen and Sperling, 1985).

It is important to note, that the models that have been discussed above are representations of the *algorithmic* structure of an elementary motion detection system. There are numerous ways how their individual computational steps could be implemented in a biological system at a biophysical, molecular or network level. Mapping algorithmic elements of motion detector models onto precise neuronal mechanisms in the brain of *Drosophila* is an ongoing research, part of which is presented in this thesis. Manuscript 1 will shed light on how asymmetric temporal filtering might be implemented via the differential temporal tunings of input neurons onto direction-selective T4 and T5 neurons. In the following chapter, I will outline anatomical and biological aspects of the visual system of the fruit fly and explain how motion processing circuits can be investigated within this model system.

1.2 VISUAL NEUROSCIENCE IN THE FLY

The great appeal of using *Drosophila melanogaster* as a model organism for motion vision research comes from an exceptionally powerful combination of experimental tools available. First, fruit flies respond in a stereotypical way to external stimuli which allows for a reliable read-out of motion-guided motor reflexes. Second, detailed circuit mapping is made possible by the manifold neurogenetic toolbox in *Drosophila* which enables targeted manipulation of neuronal mechanisms at a cellular level. Finally, the relative simplicity of the fly nervous system in contrast to the vertebrate brain makes it more feasible to identify and link neural circuits, in some cases even single neurons, to ecologically relevant behaviors. This chapter will give an overview of the state-of-the art research on motion processing in the fly and introduce the experimental techniques that have been applied in the projects presented in the main part of this thesis.

1.2.1 Motion-guided behaviour in the fly

The ultimate goal of systems neuroscience is to map circuits to functions and to explain how sensory input is processed in order to initiate an appropriate behavioral response. *Drosophila* exhibits a wide range of mostly stereotyped behaviors, a detailed study of which can help gain insight into sensory processing in the brain and often is the starting point for further efforts in circuit mapping. In this section, I will motivate research on motion processing pathways by giving a range of examples where motion detection is the foundation for the implementation of more complex and vitally important fruit fly behaviors.

Optomotor response

When confronted with full-field horizontal motion stimuli, walking or tethered flying flies try to turn with the direction of motion, a reflex that is called the *optomotor response* (Götz, 1964; Götz et al., 1979; Borst et al., 2010). The optomotor response is believed to serve as a compensatory reflex to stabilize gaze and course control during flight or walking (Götz, 1968; Borst, 2014). Because flies are lightweight animals, an external perturbation, such as a wind gust, can easily bring them off track. From the perspective of the fly, such perturbation would be reflected in an unexpected drift of its visual environment, i.e. in full-field motion, in the opposite direction. Turning in the same direction as the visual drift in order to stabilize the image on the retina therefore means counter-acting the perturbation. Thus, the optomotor response is a critical reflex to maintain stability during directed walking or flight maneuvers.

Historically, detailed observation of optomotor behavior in the snout weevil *Chlorophanus viridis*, as well as in a variety of other invertebrates, led to the proposal of the first algorithmic model of motion detection, the Hassenstein-Reichardt detector (Hassenstein, 1951; Hassenstein and Reichardt, 1956; von Fermi and Richardt, 1963; Götz, 1964). In the original study, the snout beetle was placed in a Y-maze and forced to make a decision between left

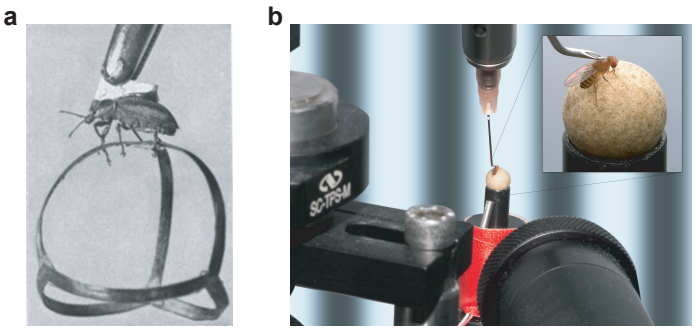


Figure 10: Experimental analysis of optomotor responses | **a**, Snout beetle walking on a Y-maze as used in early studies on optomotor behaviour (Hassenstein, 1951). **b**, Tethered walking *Drosophila* on an air-cushioned styrofoam ball, tracked by cameras and various sensors. On the computer screens surrounding the animal, arbitrary visual stimuli can be presented. This is the experimental set-up that has been used for behavioral experiments in this thesis. Photo by R. Schorner, reprinted with permission.

and right while being visually exposed to a moving background pattern (Figure 10a). With the advance of technology, in nowadays behavioral experiments with *Drosophila melanogaster* the tethered animal is placed on a treadmill and each of its movements elicited in response to external stimuli is precisely tracked and recorded (Figure 10b). This enables a more thorough and quantitative analysis of behavior. Comparison of behavioral responses following defined input stimulation with computer simulations of the fly motion vision circuitry has led to detailed insights into the mechanisms of motion detection in the fly (Borst et al., 2010). With the help of these techniques and in combination with neurogenetic manipulation of individual cell types and neurophysiological experiments, the neural substrate of motion detection in the fruit fly has been mapped to a great extent (see Section 1.2.2). Although many open questions remain, understanding of the fly motion vision circuitry has reached a point where closing the loop between sensory input and behavioral output seems at least within reach.

Collision-avoidance and landing response

While course control certainly represents one of the most important tasks for a flying animal, it is worth pointing out the diversity of other behaviors that rely on the establishment of accurate motion vision. One such example is the response to expanding objects, also called *looming stimuli*. It has been shown that presentation of such stimuli either leads to *collision-avoidance* or to *landing* responses in *Drosophila* (Tammero and Dickinson, 2002; Muijres et al., 2014). In collision-avoidance, the behavioral response of the fly can be clearly identified as an attempt to steer away from a looming object. A landing response is characterized by an extension of the legs as if the fly is approaching a surface and expects to be able to land. The contrast between the potential outcomes of these two qualitatively very different behaviors renders this choice a task quite critical for the survival of the animal. Which behavior is triggered in the experiment has been shown to depend on the exact position as well as on the velocity of the expanding stimulus (Tammero

and Dickinson, 2002). This indicates that the precise analysis of the pattern of motion cues within the visual field is a precondition for such behavioral choices. Indeed, it has been shown that collision-avoidance and landing responses in *Drosophila* depend on direction-selective cells in the fruit fly (Schilling and Borst, 2015).

Conspecific tracking during male courtship behaviour

Another example of motion-guided behaviour is the ability to identify small moving objects in cluttered natural environments. Male flies track and closely pursuit potential mating partners during their courtship ritual (Land and Collett, 1974; Hall, 1994). Attraction to visual objects has been shown to depend on the exact shape and geometry of an object and therefore certainly also involves detection of visual features other than motion (Maimon et al., 2008; Robie et al., 2010). Additionally, in the context of mating behavior, olfaction might play a role more important than vision for identification of the species and the sex of a potential mate (Dweck et al., 2015). However, it has been argued that the ability of male flies to closely follow a female fly at small distances and to chase her at high speeds involves a class of male-specific neurons which have been reported to selectively respond to small object motion in several fly species (Gilbert and Strausfeld, 1991; Nordström et al., 2006; Trischler, 2010). In Manuscript 2 of this thesis, we investigated a class of visual projection neurons in the male fruit fly which might relay visual information about the position of a female fly to the central brain by taking advantage of figure-ground discrimination through detection of relative motion (Egelhaaf, 1985).

1.2.2 Fly visual system

The fly nervous system comprises the brain and the ventral nerve cord which is located in the thorax and the abdomen of the animal. The estimated number of cells in the brain of *Drosophila* has an order of magnitude of 100,000 (Simpson, 2009). The two optic lobes, which reside laterally on each side of the brain, make up about 15,000 cells each (Figure 11a). Looking at these proportions it becomes already clear that the visual sense must be of major importance for *Drosophila* if such large amounts of its nervous infrastructure are dedicated to visual processing.

The optic lobes of the fly consist of four different neuropiles which are called *lamina*, *medulla*, *lobula* and *lobula plate* (Figure 11b). These neuropiles are structured as a hexagonally distributed array of neural cartridges which mirrors the hexagonal packing of photoreceptive units in the retina. Each neural cartridge consists of various cell-types which are so densely packed in *Drosophila* that their cell bodies are pushed to the periphery of the brain in order to create more space for the neuropile (Figure 11c). Early efforts to map all the different cell-types in the fly optic lobe have identified a range of columnar cells but also some wide-field neurons which span several columns (Fischbach and Dittrich, 1989). Each columnar cell-type is repeated per column and thus forms a retinotopic array of neurons over the whole neuropile where each neuron processes retinal information from one point in

visual space. Before introducing the relevant cell-types of the motion vision circuitry, it is worth considering how light-sensitive signals are generated in the first place in the retina.

Retina

Phototransduction – the process of converting light signals impinging on a photoreceptor into electrochemical output signals – takes place in the so-called rhabdomeres in the fly retina (reviewed in Montell (1999); Hardie and Juusola (2015)). The compound eye of *Drosophila* consists of around 800 ommatidia. In each ommatidium, light is focused through a lens onto the microvilli in the rhabdomere, which contain the light-sensitive receptor protein *rhodopsin*. Through photoisomerization, this protein undergoes a series of conformational changes upon incidence of light which ultimately transforms rhodopsin into metarhodopsin. This in turn leads to the activation of a G-protein coupled second messenger cascade and eventually to activation of the phospholipase C. Through mechanisms that still remain controversial (Montell, 1999; Hardie and Juusola, 2015), this leads to opening of the two Ca^{2+}-permeable channels TRP and TRPL. The resulting influx of cations leads to depolarization of the cell. Interestingly, while *Drosophila* photoreceptors are activated through light, phototransduction in the vertebrate retina triggers hyperpolarization of the photoreceptors, so they have sign-reversed responses (Yau and Hardie, 2009). Another distinction is that phototransduction in fly photoreceptors works extremely fast. While human cone photoreceptors cannot resolve temporal frequencies beyond 60 Hz (Hecht and Shlaer, 1936), flicker fusion frequencies between 100-200 Hz have been reported for *Drosophila* (Cosens and Spatz, 1978; Miall, 1978).

In fruit flies, each ommatidium possesses seven rhabdomeres. R1–6 express Rhodopsin 1 which is light-sensitive across most of the visual spectrum of *Drosophila* covering wavelengths from below 300 nm up to 600 nm. While R1–6 feed into the motion vision pathway, R7 and R8 give input to a separate color vision system (Yamaguchi et al., 2008). They are arranged such that R1–6 surround R7 and R8, which sit on top of each other in the center of the ommatidium. Because they are spatially displaced, the rhabdomeres receive light from slightly different visual angles according to the lens optics. However, the specific geometry of their arrangement is such that the visual angle of each rhabdomere matches exactly that of another rhabdomere in an adjacent ommatidium. Thus, there are seven rhabdomeres from seven different ommatidia that receive light from the same direction in visual space. These seven rhabdomeres then send their axons to the same neural cartridge in the lamina, which is an intriguing example of neural wiring specificity. By taking advantage of this design, called *neural superposition* (Kirschfeld, 1967), signals from seven ommatidia are pooled in one lamina cartridge leading to an increase of sensitivity but without a sacrifice in spatial resolution. The neural superposition eye has evolved in *Dipterans*, "true flies", which rely heavily on vision also under dim light conditions. Other arthropod species may be equipped with other types of compound eyes, such as the apposition eye, which requires bright daylight conditions, or the optical superposition eye, which has evolved for nocturnal animals, for example, moths.

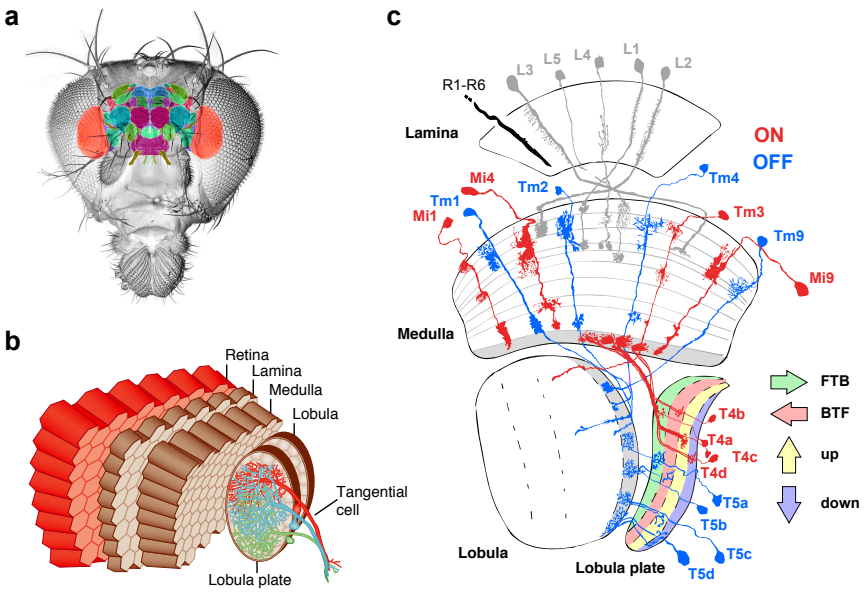


Figure 11: Fly visual system and motion vision circuitry | **a**, *Drosophila* head with a schematic of the brain. Anatomical brain structures are highlighted in different colors. Optic lobes are shown in bright red. Adapted from illustration courtesy of Kei Ito, Sheena Brown and Nicholas J. Strausfeld. **b**, Schematic of the fly optic lobe. Three lobula plate tangential cells are depicted in the lobula plate. Illustration from Borst (2014). **c**, Overview of the fly motion vision circuitry. Anatomically correct representations of relevant columnar cell types are shown in different colors. Lamina cells are colored in gray, ON-pathway cells in red and OFF-pathway cells in blue. LPi cells not shown. The four layers of the lobula plate are highlighted in color according to their preferred direction of motion (front-to-back (FTB): green, back-to-front (BTF): red, up: yellow, down: purple). Graphics adapted from Fischbach and Dittrich (1989).

Lobula plate tangential cells

The large *lobula plate tangential cells (LPTC)* (Figure 11b) are thought of as a read-out of the fly motion vision system (Borst et al., 2010). A subclass of LPTCs, the so-called *horizontal system (HS)* cells, depolarize in response to wide-field horizontal motion in front-to-back direction and hyperpolarize for the other direction. There is evidence that horizontal optomotor turning in tethered flies is controlled or triggered by HS cells (Hausen and Wehrhahn, 1983; Haikala et al., 2013; Fujiwara et al., 2017). Therefore, HS cell recordings are often shown together with behavioral responses. However, if visual motion induced by self-motion elicited responses in HS cells, the resulting optomotor response would counteract the direction of self-motion and hence force the fly to return to a straight direction of heading undermining any attempt at "voluntary" turning. If HS cells were to guide motor signals, they should therefore also implement an *efference copy*, a signal that effectively subtracts visual response contributions that are due to self-motion. Indeed, there is first evidence that HS cells not only code for visual motion but also receive suppressive signals during self-induced "voluntary" turns in tethered flight (Kim et al. (2015, 2017); but see Fujiwara et al. (2017)).

The horizontal system is complemented by the *vertical system (VS)* cells which respond more strongly to vertical than to horizontal motion. However,

closer investigation revealed that VS cells are in fact selective for complex rotational optic flow fields around various body axes which renders them ideally suited for course control during acrobatic flight maneuvers (Krapp and Hengstenberg, 1996; Franz and Krapp, 2000; Weber et al., 2010; Borst, 2014). How do these cells acquire such complex response properties?

Lobula plate – a functional map of visual motion

LPTCs possess widely arborized dendritic trees in the lobula plate where they form synaptic connections with bushy T4 and T5 cells, as well as lobula plate intrinsic (LPi) cells (Figure 11c). T4 and T5 cells as well respond to visual motion and have locally restricted receptive fields (Maisak et al., 2013). There are four subtypes of each T4 and T5 neurons, called T4a–d and T5a–d. Each subtype sends their axon terminals to one of the four layers of the lobula plate, respectively. While T4/T5-neurons arborizing in layer 1 of the lobula plate selectively respond to front-to-back motion, neurons in layer 2 respond to the opposite direction of motion (Maisak et al., 2013; Fisher et al., 2015b). The same is true for layers 3 and 4, but regarding upward and downward motion. Overall, T4 and T5 cells form a complete map for local motion cues spanning the entire visual field of the fly. The lobula plate is therefore an outstanding example for brain tissue organization into functional maps. Why does the fly need two cell types for this? It has been shown that T4 and T5 cells respond selectively only for motion of bright edges and dark edges respectively (Maisak et al., 2013). The separate processing of ON and OFF features is a general property of visual processing that seems to be shared across species (Borst and Helmstaedter, 2015).

LPi neurons are inhibitory cells that locally pool T4 and T5 signals and inhibit LPTCs when a stimulus moves in the null direction of the cell (Mauss et al., 2014, 2015). They correspond to the subtractive stage in a fully symmetrical Hassenstein-Reichardt motion detector (see Section 1.1.4 and Figure 8) (Borst et al., 2010). Overall, such fully symmetrical correlation-type motion detectors account exceptionally well for a wide range of experimentally observed LPTC response characteristics (Borst and Egelhaaf, 1989; Borst et al., 2010). Complex rotational flow field selectivity, as described in VS cells, can be explained by the anatomy of their dendritic trees which selectively pick up signals from excitatory and inhibitory locally direction selective units, from T4/T5 and LPi cells respectively, in an appropriate way across all four lobula plate layers (Hopp et al., 2014).

However, measuring signals from LPTCs cannot reveal how direction selectivity emerges in T4 and T5 cells in the first place. This is because the subtractive stage in a full Hassenstein-Reichardt detector yields a difference signal between opposed motion half-detectors, which prohibits a read-out of the absolute signal amplitudes in the half-detectors (Haag et al., 2016). Whether direction selectivity is due to *preferred direction enhancement* or *null direction suppression* can therefore not be assessed at the level of LPTCs. The emergence of direction selectivity can be pinpointed to the dendrites of T4 and T5 cells which thus represent the non-linear interaction stage of a motion half-detector (see Figure 8) (Maisak et al., 2013; Fisher et al., 2015b).

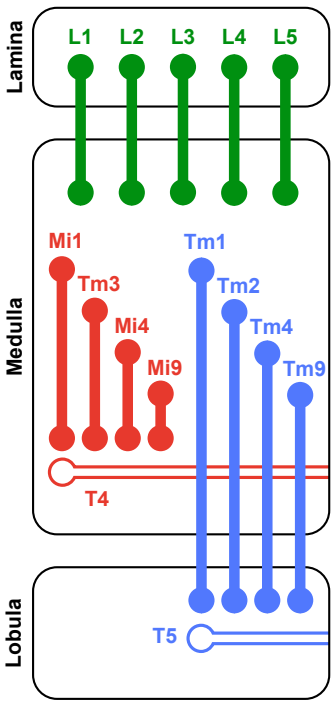


Figure 12: Schematic of the fly motion vision pathway | Lamina neurons L1–5 (green) provide input signals to the fly motion vision circuitry. Signals are then split into an ON- and an OFF-pathway of motion vision. In the medulla, the ON-pathway comprises Mi1, Tm3, Mi4 and Mi9 cells (red). In the OFF-pathway, Tm1, Tm2, Tm4 and Tm9 cells (blue) are found in the medulla. Finally, direction selectivity emerges in T4 (red) and T5 (blue) cells which respond to bright or dark moving edges, respectively.

How then do T4 and T5 cells become direction selective in the first place? Before addressing this question, it is instructive to provide a short overview of synaptic partners and the motion vision circuitry presynaptic to T4 and T5.

Motion vision pathway in lamina and medulla

In the lamina, large monopolar cells L1, L2 and L3 pick up signals from R1–6 via histaminergic synapses (Figure 11c, Figure 12) (Hardie, 1989). Although the lamina consists of an intricate network involving monopolar cells, photoreceptor terminals, amacrine cells and other neurons, I will focus here only on the most important connections that presumably give direct input to motion processing channels. There are two other monopolar cells, L4 and L5 which receive input from L2 and L1, although L4 also makes direct connections with R6 (Rivera-Alba et al., 2011; Takemura et al., 2013). L1 and L2 are the two main inputs to the ON- and OFF-pathway of motion vision, respectively (Joesch et al., 2010; Clark et al., 2011). However, both neurons respond transiently to OFF stimuli (Reiff et al., 2010; Clark et al., 2011; Freifeld et al., 2013). This is readily explained by electrical coupling of L1 and L2 through gap junctions (Joesch et al., 2010) and the OFF-polarity of the presynaptic photoreceptors in *Drosophila* (see above). The ON-OFF split only emerges at their synapses: While L2 is purely cholinergic and forms sign-conserving synapses with its postsynaptic partners in the

medulla (Takemura et al., 2011), L1 releases glutamate which can act as an inhibitory neurotransmitter in invertebrates via glutamate-gated chloride channels (Liu and Wilson, 2013; Mauss et al., 2014).

L1's main postsynaptic targets in the medulla are the columnar neurons Mi1 and Tm3 (Takemura et al., 2013). Recordings from these neurons suggest that the L1 synapse must indeed be sign-inverting, since both Mi1 and Tm3 depolarize transiently to global and local ON flicker stimuli (Strother et al., 2014; Behnia et al., 2014; Yang et al., 2016). With respect to the number of synapses, Mi1 and Tm3 are major synaptic input partners to direction selective T4 cells (Takemura et al., 2013).

In the OFF-pathway, L2 axon terminals form synaptic contact with Tm1, Tm2 and Tm4 cells in the medulla (Figure 11c, Figure 12) (Takemura et al., 2011, 2013). All three cells respond with transient depolarization when presented with OFF stimuli (Meier et al., 2014; Behnia et al., 2014; Serbe et al., 2016; Yang et al., 2016). Ultrastructural investigation using electron microscopy found that all three cells provide strong input to T5 neurons, as does also the medulla neuron Tm9 (Shinomiya et al., 2014). In contrast to the other three OFF-pathway neurons, Tm9 has tonic response properties and exhibits sustained depolarization upon presentation of OFF-flicker stimuli (Fisher et al., 2015a; Serbe et al., 2016). This property is most likely inherited from its main input partner L3 (Silies et al., 2013). Controversially, while Serbe et al. (2016) found Tm1, Tm2, Tm4 and Tm9 to possess local receptive fields and be suppressed by wide-field stimuli, Tm9 has been described as a wide-field neuron in a different study (Fisher et al., 2015a).

All aforementioned input neurons to T4 (Mi1 and Tm3) and T5 (Tm1, Tm2, Tm4 and Tm9) have been reported to respond independently of the direction of a moving stimulus (Behnia et al., 2014; Meier et al., 2014; Serbe et al., 2016). Calcium imaging experiments revealed that the first direction selective stage in this cascade of visual signal processing are the dendrites of T4 and T5 cells (Maisak et al., 2013; Fisher et al., 2015b).

Only recently, new electron-microscopic circuit reconstructions have indicated that two other medulla neurons, Mi4 and Mi9, synapse onto T4 dendrites (Louis Scheffer, Janelia Research Campus, personal communication). It will be interesting to investigate the functional properties of those new cell types and their relevance for motion detection in the ON-pathway.

Elementary motion detection in T4/T5 neurons

Basic models of motion detection require at least two spatially offset input channels and an asymmetry regarding their temporal filtering properties. Direction selectivity then emerges from non-linear interactions between the input channels either leading to *preferred direction (PD) enhancement* in the Hassenstein-Reichardt (HR) detector, or to *null direction (ND) suppression* in the Barlow-Levick (BL) detector (see Figure 8).

One way to experimentally distinguish between those two models of motion detection are apparent motion stimuli. Here, a moving stimulus is discretized into a sequence of adjacent bars or squares. Presentation of these in the correct temporal sequence resembles a motion stimulus and can therefore elicit direction selective responses. On the other hand,

inordinate flashing of the individual bars or squares in a randomized fashion does not trigger direction selective responses and permits assessment of a linearly expected response as the sum of the individual flicker responses. If the response to the apparent motion stimulus in PD is higher than the linear expectation, the interaction is *supra-linear*, which is a hallmark of PD enhancement. If the linear expectation is larger than the response to ND apparent motion, then the response is *sub-linear*, which is indicative of ND suppression.

In T4 cells, a supra-linear response component could be detected using apparent motion stimuli in calcium imaging (Fisher et al., 2015b). It has been argued that Mi1 and Tm3 implement the two input channels of an HR detector based on a small timing difference between their impulse responses as measured in whole-cell recordings (Behnia et al., 2014). T4 dendrites indeed receive spatially offset input from Mi1 and Tm3 cells (Takemura et al., 2013). However, their spatial arrangement does not match their timing difference in an HR-like scheme of motion detection. Additionally, the reported spatial and temporal offsets between Mi1 and Tm3 are both so small that even mild deviations from the reported average values would cause the motion detector to break down. The hypothesis that Mi1 and Tm3 constitute the only and defining input channels to an HR detector has been rejected by means of blocking experiments which showed that ON motion responses in the fly are abolished when blocking Mi1 but remain partly intact when blocking Tm3 cells (Ammer et al., 2015).

Similarly in the OFF-pathway, an HR detector was proposed based on a time lag between Tm1 and Tm2 and supra-linear response properties in T5 dendrites (Behnia et al., 2014; Fisher et al., 2015b). Also here, the latency between Tm1 and Tm2 was found to amount to only 13 ms, which is rather small (Behnia et al., 2014). Electron microscopy based circuit reconstructions confirmed a corresponding displacement between Tm1 and Tm2 synapses on the T5 dendrite as well as between Tm9 and Tm2 (Shinomiya et al., 2014). In fact, Tm9 and Tm2 would be ideally suited to implement asymmetric temporal filtering in an HR detector due to their vastly different temporal filter characteristics (Serbe et al., 2016). However, blocking experiments again failed to yield a clear resolution of the question since neither single-cell blocks of Tm1, Tm2, Tm4 or Tm9 nor combinatorial blocks could completely impair responses in HS cells or motion guided behaviour (Serbe et al., 2016). Therefore, it seems likely that more elaborate computations involving interactions between more than two input channels take place on either T4 and T5 cell dendrites.

A recent study measured calcium responses in T4 cells to apparent motion stimuli using a new telescopic stimulation device that can target light stimuli precisely onto the neural cartridges in a neural superposition eye (Haag et al., 2016). By varying the position of the apparent motion stimulus within the T4 receptive field, Haag et al. (2016) found evidence for both an enhancing but also a suppressive non-linearity. In fact, there seem to exist two spatially segregated regimes of PD enhancement and ND suppression along the receptive field of a T4 cell. These findings led to the proposal of the three-arm hybrid detector that combines both types of non-linearity and thus achieves higher direction selectivity (see Figure 9). Another study used

white noise stimuli to map the linear receptive field structure of T4 and T5 cells and found spatiotemporally tilted receptive fields (Leong et al., 2016). Based on this, the study argues as well that direction selectivity emerges from a combination of both mechanisms in both ON- and OFF-pathways of motion vision. Finally, a third study went one step further and mapped not only first-order linear but also second-order receptive fields of T4 and T5 (Salazar-Gatzimas et al., 2016). They verified their results using a newly developed correlated noise stimulus to obtain T4 and T5 calcium responses for different types of pairwise pixel correlations. For both cell types, they argue for a linear-nonlinear model of motion detection which resembles an HR detector with a static nonlinearity instead of a multiplication. However, they do not consider the possibility of a hybrid detector in their publication. While PD enhancement is clearly more pronounced in their experiments, some of their observations could be due to ND suppression and would certainly have to be re-evaluated with respect to the new hybrid detector scheme. Overall, there is increasing evidence that the direction selectivity of T4 and T5 dendrites could indeed emerge through a combination of both mechanisms for both cell types. Future studies will have to address these questions in further detail and involve additional experimental tools like electrophysiology or voltage imaging.

1.2.3 *Drosophila* neurogenetic toolbox

History of research on *Drosophila melanogaster* dates back to more than 100 years ago (Morgan, 1910; Bellen et al., 2010). In this time span, an enormous amount of knowledge about the fruit fly has been collected in the fields of developmental biology, neuroscience, genetics and more. Specifically in genetics research, *Drosophila* has a long historical record as a model organism. The resulting advance of genetic engineering methods for the fruit fly has boosted research progress in other areas, one of them being neuroscience. With modern techniques it is feasible to generate thousands of genetically modified fly strains, each of them with a different insertion that can be used to target subsets of neurons throughout the whole brain of *Drosophila* (Pfeiffer et al., 2008). This facilitates fly neuroscience in unprecedented ways: nowadays, *Drosophila* neuroscientists have large "libraries" of modified transgenic fly strains at their disposal, each one with unique insertions, that only need to be crossed in appropriate ways in order to drive expression of a multitude of available effector tools in a target cell population. The short life cycle of *Drosophila* and the ease of culturing and maintaining large numbers of fly strains do their part to shorten the time span between hypothesis and experiment, which greatly contributes to the efficacy of *Drosophila* neuroscience.

Targeted gene expression

Manipulation of individual neurons is a prerequisite for neural circuit analysis. In order to target expression of arbitrary genes to specific neurons in *Drosophila*, binary expression systems such as the Gal4/UAS or the LexA/lexAop system have been developed (Venken et al., 2011). Binary

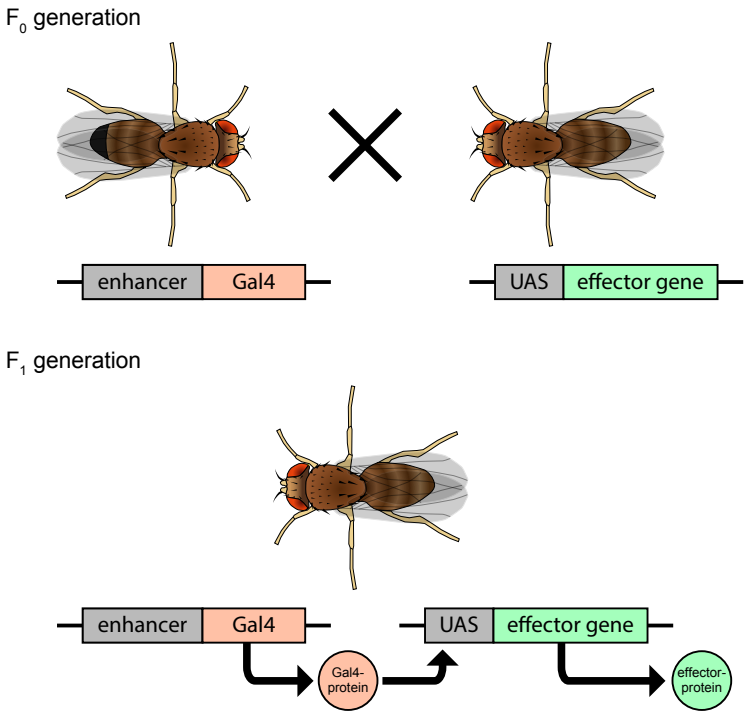


Figure 13: Crossing scheme for the Gal4-UAS system | In the parental F_0 generation, a driver fly line is crossed to an effector line (top). Expression of the effector protein is triggered in F_1 generation flies through binding of the Gal4 protein to UAS (bottom).

expression systems rely on patterned expression of a *driver* gene which activates expression of an *effector* gene, that can be any gene of interest. Depending on the insertion site in the genome, the expression pattern of the driver is controlled by endogenous enhancers that determine the neural expression pattern of the driver line. In other words, the driver line dictates the *where* while the effector line specifies *what* will be expressed.

In the Gal4/UAS system, the yeast transcription factor Gal4 binds to a specific sequence of 17 basepairs which is called the *Upstream Activating Sequence (UAS)* (Brand and Perrimon, 1993). UAS controls expression of the neighboring effector gene. Because the exogenous Gal4 protein lacks any endogenous targets in *Drosophila*, it will only activate sequences under the control of UAS and not interfere with endogenous gene expression. The expression of the effector is therefore confined to the expression pattern of the Gal4 protein.

In practice, targeted gene expression using the Gal4/UAS-system can easily be accomplished by crossing an appropriate driver fly line with a specific effector line (Figure 13). The subsequent generation of flies inherits both, the driver as well as the effector gene, which triggers selective expression of the chosen effector protein.

Large numbers of Gal4 lines have been created by random insertions through injection of a vector containing transposable P-elements into the embryo of a fly, e.g. by Hayashi et al. (2002). The lack of control with this method over the specific insertion site, however, often results in relatively broad expression patterns including several types of neurons. This therefore renders targeted manipulation of specific neurons challenging. To overcome

these limitations, site-specific integration of plasmids expressing the Gal4 protein using the Φ31 integrase is achieved in more recent approaches to transgenesis of driver lines (Venken et al., 2011). Using these methods, expression levels as well as the specificity of driver lines could be improved and libraries comprising several thousands of fly strains have been generated for neuroscientific purposes (Pfeiffer et al., 2008; Jenett et al., 2012).

Still, binary drivers often lack the degree of specificity which would be essential for the conduct of clear and conclusive experiments when manipulating neural circuits. For this reason, intersectional strategies like the split-Gal4 system have been developed (Luan et al., 2006). Here, the Gal4 protein is split into an AD and a DBD domain which are expressed independently from each other by using different driver lines. Neither domain is functional on its own. However, transcription of the effector gene through Gal4 will be activated exclusively in those cells where the expression patterns of both driver lines overlap and a functional Gal4 protein can be reconstituted by fusion of the two domains. In another intersectional strategy which is frequently used, the expression of Gal4 is suppressed by independent expression of the protein Gal80 (Lee and Luo, 1999).

With these tools available, targeted manipulation of selected cell types is feasible for most neurons in the *Drosophila* motion vision pathway and expression of distinct effectors enables precise circuit mapping and functional analysis. In the following, I will introduce a few of the most common effector proteins, with focus on those that have been used in my doctoral projects.

Visualization

Targeted expression of a fluorescent dye in a neuron makes its anatomy visible under a fluorescence microscope. In 2008, Osamu Shimomura, Martin Chalfie and Roger Tsien received the Nobel Prize in Chemistry for the discovery and development of the *green fluorescent protein (GFP)* (Chalfie et al., 1994). GFP is a protein which is endogenously expressed in jellyfish and that emits green light of about 500 nm wavelength when activated by shorter wavelength excitation light close to the UV spectrum. Since its discovery, GFP has been used extensively in neuroscience and other areas as a fluorescent marker for cell visualization. Apart from GFP, fluorescent markers with emission spectra at many different wavelengths have been developed and are constantly subject to improvement through genetic engineering (Shaner et al., 2004). Labelling neurons with transgenic fluorescent dyes allows for visualization of structural and anatomical details.

Functional imaging

Another advantage of labelling neurons with fluorescent proteins is that it makes non-invasive imaging of the targeted cell population *in vivo* possible. *Genetically encoded calcium indicators (GECI)* have been developed to take full advantage of this and enable optical read-out of neuronal activity (Mank and Griesbeck, 2008). GECIs are mutants of auto-fluorescent proteins, such as GFP, that have been modified so as to change their fluorescent properties depending on the concentration of Ca^{2+} ions in the cell. Depolarization of a neuron causes voltage-gated calcium channels in the cell membrane to

open. The resulting influx of Ca^{2+} ions increases the intracellular calcium concentration which can be read out through GECIs. Calcium is hence coupled to membrane voltage and can therefore be used as a proxy for neuronal activity.

For the calcium indicator GCaMP, the GFP protein has been fused to an M13 peptide and the calcium-binding protein calmodulin. Binding of calcium to the calmodulin domain induces conformational changes in this construct which in turn promotes deprotonation of the chromophore and thus leads to brighter fluorescence. There is a couple of limitations one should be aware of when using GCaMP or other GECIs as a proxy for membrane voltage: First, a GECI binds to calcium and therefore also acts as a calcium buffer, which changes the concentration of free cytosolic calcium. If the concentration of calcium indicator is too high, it might significantly alter calcium dynamics in the cell and nonlinearly perturb the system (Borst and Abarbanel, 2007). Second, calcium channels inactivate when the cell hyperpolarizes which effectively leads to half-wave rectification in the voltage-to-calcium transformation. Third, calcium binding to the indicator protein has limited forward and backward rate constants which reduces the temporal precision of the signal. Even for the presumably fastest version, GCaMP6f, decay time constants between 200–400 ms have been reported (Chen et al., 2013).

In Manuscript 3 of this thesis, we try to overcome the problem of slow calcium indicator dynamics by using the glutamate sensor iGluSnFR which reports synaptic activity by changing its fluorescence depending on the glutamate concentration (Marvin et al., 2013). The decay rate of the iGluSnFR protein is a lot faster than that of GCaMP and therefore it induces fewer temporal distortions. In addition, it might be beneficial to measure synaptic release of a neurotransmitter directly rather than calcium concentration or membrane voltage. Ultimately, the amount of neurotransmitter released into the synapse is the actual signal that is received by postsynaptic neurons. However, this technique is obviously limited to glutamatergic neurons, e.g. L1 neurons in the fly motion vision circuitry.

Efforts are also going into the design of voltage indicators that can report the membrane voltage of a cell directly. Such proteins usually consist of a transmembrane voltage sensing domain fused to a reporter fluorescent protein (Lin and Schnitzer, 2016). Although voltage indicators have been successfully applied in the fly, their signal-to-noise ratio is still relatively low so that many repetitions of the same stimulus are necessary in order to obtain clear results (Yang et al., 2016). Additionally, voltage indicators still suffer from fast photobleaching, a phenomenon that affects all fluorescent dyes but is in practice much less severe for, e.g., GCaMP6 (Chen et al., 2013).

Activation tools

In order to interrogate a circuit, activation or silencing of selected neurons might be necessary. Several effector tools have been developed for this purpose. Neuronal activation can for example be achieved by expressing the temperature gated cation channel dTRPA1 in the cells of interest (Berni et al., 2010). dTRPA1 is a channel which is innately involved in thermo-

tactic behavior in *Drosophila* due to its temperature sensitive conductance (Rosenzweig et al., 2005; Hamada et al., 2008). Temperatures above 25° lead to opening of the channel and thus strong depolarization of the neuron. This is reversible by decreasing the temperature again. Artificially driving expression of dTRPA1 via binary expression systems therefore gives the experimenter a tool that allows them to control membrane voltage and activate specific neurons by setting the temperature. Temperature control, however, is a relatively slow process which is why dTRPA1 channels are not well suited for instantaneous activation of neurons.

Optogenetic activation tools promise higher temporal precision instead. For example, the ATP-gated ionotropic purinoceptor P2X2 can be activated via photo-stimulated release of caged ATP (Lima and Miesenböck, 2005). Even more potential lies in intrinsically light-activatable channel proteins. Channelrhodopsin is a light-gated ion channel that is innately involved in phototaxis behaviour of a species of algae (Harz and Hegemann, 1991). Transgenic efforts have successfully established this protein as a photo-activatable on-switch for neurons (Boyden et al., 2005; Nagel et al., 2005). Channelrhodopsins are typically excitable with blue light in a spectrum of wavelengths of approximately 450–490 nm. Such short wavelengths are usually limited in their penetration depth of neural tissue, which is why invasive methods such as light delivery via implanted light guides often remain necessary. In the visual system of *Drosophila* specifically, there is the additional caveat for possible cross-talk between optogenetic stimulation and the visual system of the fly due to the close distance of the fly photoreceptors which are sensitive to light in this spectrum. Therefore, red-shifted variants of the channelrhodopsin protein have been generated, such as the red-activatable channelrhodopsin ReaChR (Lin et al., 2013) or the even farther red-shifted Chrimson (Klapoetke et al., 2014). Such red-sensitive agents have deeper penetration depths and in *Drosophila*, they allow for non-invasive delivery of optogenetic stimuli through the cuticle while minimizing interference with the fly visual system (Bath et al., 2014). With these tools, millisecond-timescale control of neuronal activity through optical triggers becomes possible. While spatial control over the elicited pattern of neural activation is mostly still carried out by genetic means through targeted expression of the optogenetic agent, new stimulation strategies are being explored with respect to shaping the light path so as to gain precise spatial control over the delivery of the optogenetic stimulus, e.g. by employing holographic techniques (Reutsky-Gefen et al., 2013).

Silencing tools

For silencing experiments, two main strategies are worth considering. One way is to block synaptic output by using effectors which interfere with synaptic mechanisms. One such agent is shibire^{ts1}, a temperature-sensitive allele of the protein dynamin which is essential for recycling of synaptic vesicles in *Drosophila* (Kosaka and Ikeda, 1983; Kitamoto, 2001). Close to room temperature, the protein functions normally but it becomes defective at higher temperatures (>29°). In the latter case, it leads to gradual ($\approx$ 1 min) depletion of synaptic vesicles and thus to inhibition of synaptic output.

Decreasing the temperature reverses this effect again. Thus, shibire^{ts1} together with genetically targeted expression gives researchers the possibility to silence selected neurons with mild temporal precision. Another way to block synaptic release is the tetanus toxin light chain TNT (Sweeney et al., 1995). Tetanus toxin is a powerful neurotoxin which cleaves the synaptic protein synaptobrevin that is involved in vesicle release. Expressing TNT in a neuron therefore abolishes synaptic transmission, an effect which is not reversible. However, TNT does not interfere with the electrophysiology of a cell, which is why it is still possible to measure voltage or calcium signals from TNT-targeted neurons.

Instead of disrupting synaptic mechanisms, "electrical" silencing of neurons can be achieved by hyperpolarizing them. Overexpression of the inwardly rectifying potassium channel Kir2.1 constantly hyperpolarizes targeted neurons and thus suppresses their excitability (Johns et al., 1999; Baines et al., 2001).

Similarly to the methods mentioned above, also here the development of optogenetics has endowed neuroscientists with new effector tools that can be optically triggered and thus enable neuronal silencing with high temporal precision. The *Natromonomas pharaonis* halorhodopsin is a light-gated chloride ion pump which triggers hyperpolarization when overexpressed in a neuron and activated with light near 590 nm wavelength (Zhang et al., 2007). Recently, more sensitive effectors, like the light-gated anion channels GtACR1 and GtACR2, have been derived from the cryptophyte alga *Guillardia theta* (Govorunova et al., 2015). Combination of different optogenetic tools with distinct excitation spectra and opposing effects on the membrane voltage can be used for multimodal all-optical control over activation and silencing experiments (Zhang et al., 2007).

1.2.4 Physiological techniques

Neurons encode and process information in the form of electrical signals. Membrane potential is the difference in the electric potential between inside and outside of a neuron. A minute balance of the concentration gradients of different ions, amongst others involving sodium, potassium and chloride, establishes an electrical gradient which can be controlled through opening and closing of specific ion channels. But how can we measure these signals *in vivo*? This section summarizes some of the experimental methods for assessment of the membrane voltage and live monitoring of neuronal signals.

Electrophysiology

Electrophysiological methods measure electrical activity of a neuron directly by placing electrodes into the neural tissue. Extracellular recordings are used to detect spiking activity from outside of a cell. This is possible because an electrode which is brought only in the vicinity of a spiking neuron is still susceptible to the strong depolarization during an action potential, even though signal amplitude might be reduced depending on distance. In the fly visual system, however, not all neurons show spiking activity and

a lot of information is conveyed by graded changes of membrane potential (Haag and Borst, 1998). In order to measure membrane voltage more directly, intracellular recordings are necessary. Sharp electrode recordings have been used for functional characterization of HS and VS cells in the lobula plate of large flies like *Calliphora* (Hausen, 1976; Krapp et al., 1998). Here, neurons are penetrated using sharp micropipettes with a fine pore at the tip. Through the pore the intracellular fluid is in contact with the solution inside of the pipette where an electrode is placed. This enables measurement of transmembrane voltage with high precision. Sharp electrode recordings, however, are hard to obtain for small neurons of the size as we find in *Drosophila*. Patch-clamp recordings have turned out be the more successful approach here (Sakmann and Neher, 1984). In order to patch-clamp a neuron, a slightly larger microelectrode is placed next to the cell body and by gentle suction a patch of the cell membrane is drawn into the tip of the electrode. This leads to formation of a so-called *giga-ohm seal* between the cell membrane and the tip of the glass pipette. Upon application of more suction, the membrane patch in the pipette can be detached. Because the electrode stays sealed to the rest of the cell body, this leads to very stable low-impedance electrical access to the cell interior. Recording from LPTCs in *Drosophila* have established fundamental insights into fly motion processing using *in vivo* patch-clamp recordings (Joesch et al., 2008).

Two-photon imaging

Electrophysiology measures the relevant quantity, membrane voltage, directly and with high temporal precision but has the great disadvantage that it is an invasive method. An additional drawback to intracellular recordings is that they are usually restricted to one neuron, although double patch-clamp recordings might be feasible (Haag and Borst, 2001). Finally, while it is possible to record from LPTCs in the *Drosophila* visual system, establishing electrophysiological access is extremely challenging for even smaller neurons such as the ones found in the lamina or the medulla. Moreover, neuronal cell bodies in *Drosophila* reside at the periphery of the brain which means that somatic signals might not be a faithful representation of the true dendritic or axonal signals due to the long cable length between soma and axons or dendrites. Therefore, in *Drosophila* neuroscience, calcium imaging is the method of choice in many cases. Imaging is minimally invasive and permits access to sub-cellular structures such as dendrites or axons and simultaneous recordings of several targets. How are calcium signals measured?

Fluorescence requires excitation by light with an appropriate wavelength. For GFP-derived GECIs like GCaMP, the peak of the emission spectrum lies at a wavelength of around 520 nm while absorption is maximal for slightly shorter wavelengths of approximately 480 nm. At this wavelength, penetration depth into neural tissue is not very large and visual artifacts due to photoreceptor stimulation can be triggered by the excitation light beam. Fluorescence relies on excitation of the fluorophore by an impinging photon containing enough energy to elevate the molecule into an excited singlet or triplet state. Relaxation of the fluorophore then releases that energy in form

of a photon with longer wavelength. Crucially, the fluorophore can also be excited by two photons carrying half of the energy each (Göppert-Mayer, 1931). Since half of the energy implies twice the wavelength, this means that it is possible to use near-infrared light of about 920 nm wavelength to excite GFP or GCaMP. Infrared light penetrates deeper into the tissue and has the advantage that it does not interfere with the visual system. For these reasons, combined confocal laser scanning microscopy and two-photon excitation has become the standard in the field for calcium imaging experiments (Denk et al., 1990).

1.3 CONCLUDING REMARKS

Although the motion vision system in *Drosophila* has been described in great detail, it still remains elusive how exactly direction selectivity emerges on T4/T5-dendrites. Before I started my doctoral work, functional descriptions of the known synaptic input elements to T4 and T5 cells had been made publicly available and several studies argued for the implementation of an HR-like scheme of motion detection. However, a number of blocking experiments unveiled conflicting evidence regarding that model. Recently, several studies have concluded that a more complex implementation of elementary motion detection might resolve these questions, moving the three-arm hybrid detector model into the focus of attention. In order to address the question of how such a model could be implemented in the fruit fly, it is necessary to know the functional processing properties of all neurons which are possibly involved. However, functional characterization of T4/T5-input elements has been performed in different studies using various methods and for distinct subsets of neurons. My first project was therefore a more systematic approach to comprehensively map the processing properties of all input elements including the cell types Mi4 and Mi9, which had not been functionally described before.

In the course of my doctoral work, more questions related to motion processing in the fly came up: Which neurotransmitters are involved in the computations of this system and how do our functional descriptions relate to neurotransmitter signaling? But also questions regarding the subsequent processing of motion information: How can the fly take advantage of its elaborate motion vision system to detect and track moving visual objects in cluttered natural environments? Finally, we asked how motion vision can be robust at all in the face of the immense amount of variability that is pervasive in natural environments. While our models of motion detection work reasonably well for artificial stimuli and capture many features of motion-guided behavioral responses in the fly, they tend to have unstable performance when stimulated using more naturalistic stimuli.

The findings that me and my collaborators made were published in three peer reviewed articles and yielded one additional manuscript that is currently submitted to a peer-reviewed journal. In the following chapter, all manuscripts are presented in a chronological order.

2 PUBLICATIONS

2.1 THE TEMPORAL TUNING OF THE *Drosophila* MOTION DETECTORS IS DETERMINED BY THE DYNAMICS OF THEIR INPUT ELEMENTS

SUMMARY This study mapped the spatiotemporal receptive fields of Mi1, Tm3, Mi4, Mi9 in the ON-pathway and Tm1, Tm2, Tm3 and Tm9 in the OFF-pathway using stochastic stimuli and reverse correlation.

Input neurons to T4 and T5 had been functionally described before, but not all of them in a single and comprehensive study. The response properties of T4 input neurons Mi4 and Mi9 had not been measured before. Using two-photon imaging, we recorded calcium responses of these neurons to white noise stimuli and performed reverse correlation analysis to obtain their spatiotemporal receptive fields. While Mi1, Tm3, Tm1, Tm2 and Tm4 showed temporal band-pass characteristics, Mi4, Mi9 and Tm9 had low-pass filter characteristics. In the spatial domain, band-pass cells, with the exception of Tm3, showed moderate surround antagonism. Low-pass filter cells Mi4, Mi9 and Tm9, however, had a strong antagonistic surround. All neurons exhibited faster response kinetics upon pharmacological application of chlordimeform, an octopamine agonist which has been used before to mimic behavioral state modulation in flies. Finally, we assessed the direction selectivity of a three-arm motion detector model with different configurations of input cell on the arms of the detector. Our simulations suggested that low-pass filter cells Mi4, Mi9 and Tm9 should be placed on the lateral arms of a three-arm detector to maximize the direction selectivity of the model.

This article was published in *Current Biology* in April 2017 (Arenz et al., 2017).

AUTHORS Alexander Arenz, **Michael S. Drews (co-first author)**, Florian G. Richter, Georg Ammer and Alexander Borst

CONTRIBUTIONS A.A., **M.S.D.**, and A.B. conceived the study and designed the experiments. A.A. conducted and analyzed the measurements of T4/T5 cell responses. **M.S.D.** designed the projector-based stimulation arena and performed and analyzed the measurements of the OFF-pathway elements. F.G.R. performed and analyzed the experiments describing the ON-pathway neurons. G.A. performed and analyzed the patch-clamp recordings from lobula plate tangential cells. **M.S.D.** performed the computer simulations. A.A. wrote the manuscript with the help of all authors.

Article

Current Biology

The Temporal Tuning of the *Drosophila* Motion Detectors Is Determined by the Dynamics of Their Input Elements

Highlights

- Input neurons to the fly motion detectors show diverse temporal filter properties
- Octopamine activation accelerates detector velocity tuning and input cell dynamics
- Dynamics of input neurons correctly predict velocity tuning of motion detectors

Authors

Alexander Arenz, Michael S. Drews, Florian G. Richter, Georg Ammer, Alexander Borst

Correspondence

arenz@neuro.mpg.de

In Brief

Recent algorithmic models of visual motion detection in the fly rely on processing of three input channels. By characterizing the spatiotemporal response properties of the neurons presynaptic to fly motion-sensing cells across two tuning states, Arenz, Drews, et al. predict functional roles for these neurons in computing the direction of motion.

Arenz et al., 2017, Current Biology *27*, 929–944
April 3, 2017
http://dx.doi.org/10.1016/j.cub.2017.01.051

CellPress

Current Biology
Article

The Temporal Tuning of the *Drosophila* Motion Detectors Is Determined by the Dynamics of Their Input Elements

Alexander Arenz,[1,2,3,*] Michael S. Drews,[1,2] Florian G. Richter,[1] Georg Ammer,[1] and Alexander Borst[1]
[1]Department of Circuits—Computation—Models, Max Planck Institute of Neurobiology, Am Klopferspitz 18, 82152 Martinsried, Germany
[2]Co-first author
[3]Lead Contact
*Correspondence: arenz@neuro.mpg.de
http://dx.doi.org/10.1016/j.cub.2017.01.051

SUMMARY

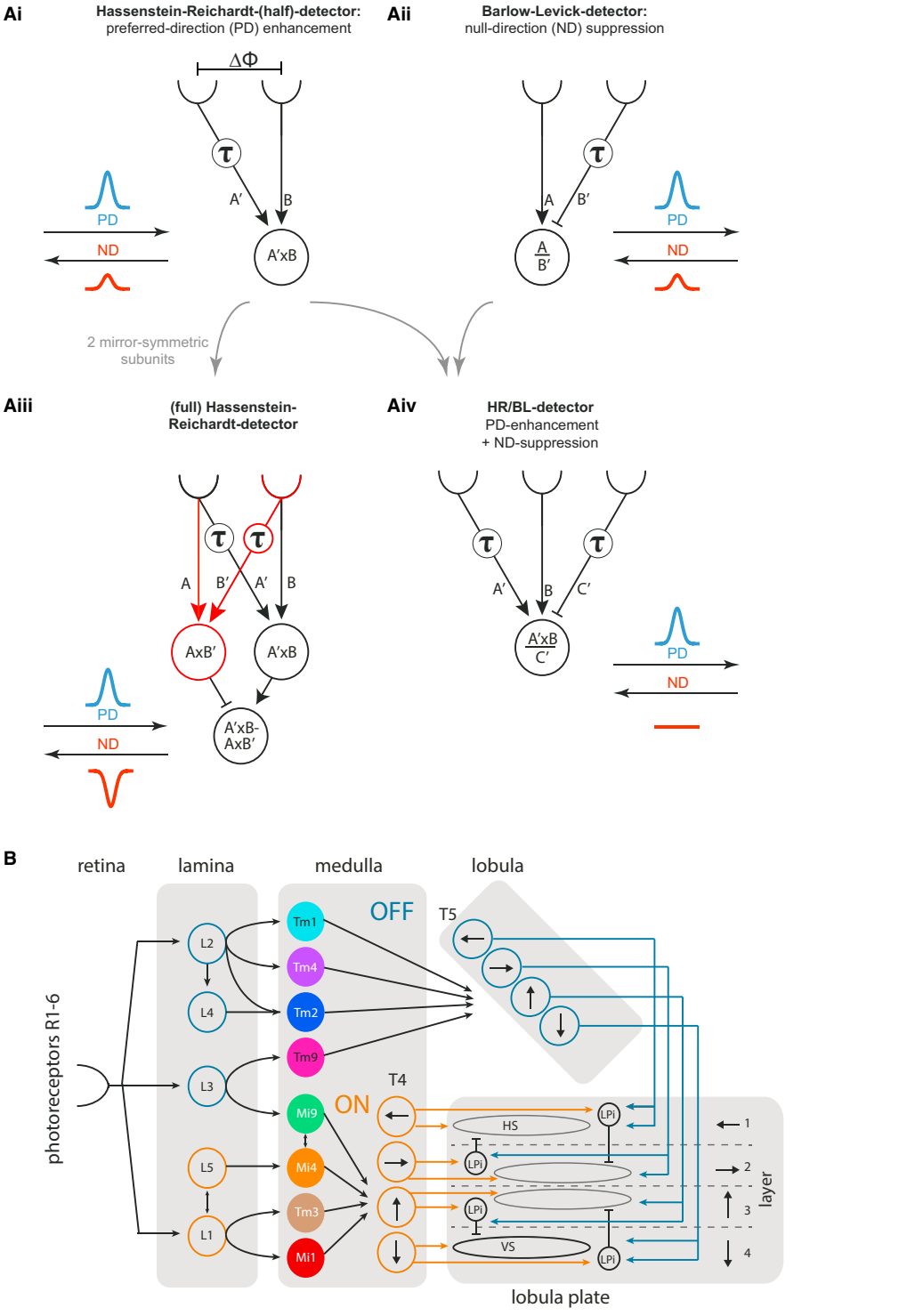
Detecting the direction of motion contained in the visual scene is crucial for many behaviors. However, because single photoreceptors only signal local luminance changes, motion detection requires a comparison of signals from neighboring photoreceptors across time in downstream neuronal circuits. For signals to coincide on readout neurons that thus become motion and direction selective, different input lines need to be delayed with respect to each other. Classical models of motion detection rely on non-linear interactions between two inputs after different temporal filtering. However, recent studies have suggested the requirement for at least three, not only two, input signals. Here, we comprehensively characterize the spatiotemporal response properties of all columnar input elements to the elementary motion detectors in the fruit fly, T4 and T5 cells, via two-photon calcium imaging. Between these input neurons, we find large differences in temporal dynamics. Based on this, computer simulations show that only a small subset of possible arrangements of these input elements maps onto a recently proposed algorithmic three-input model in a way that generates a highly direction-selective motion detector, suggesting plausible network architectures. Moreover, modulating the motion detection system by octopamine-receptor activation, we find the temporal tuning of T4 and T5 cells to be shifted toward higher frequencies, and this shift can be fully explained by the concomitant speeding of the input elements.

INTRODUCTION

The detection of visual motion arising from ego-motion is crucial for course stabilization in flies [1]. Sets of large tangential cells in the lobula plate of the fly optic lobe respond selectively to the optic flow resulting from whole-body rotation around different axes. As single photoreceptors respond to local luminance changes in a non-direction-selective way, the intervening circuitry of the optic lobe [2–5] (Figure 1) must serve to extract the feature of visual motion by spatiotemporal comparison of the responses of neighboring photoreceptors.

Two competing algorithmic models of motion detectors have been proposed (Figure 1A). Both models rely on asymmetric temporal filtering of two input signals that are then fed into a non-linearity. They differ by the type of non-linearity employed and the location of the delay filter. In the Barlow-Levick (BL) detector (Figure 1Aii) [6], the delay is located on the preferred side and the non-linearity is inhibitory, leading to a suppression of signals moving in the null direction (ND). In the Hassenstein-Reichardt (HR) detector (Figure 1Ai) [7], the delay is located on the null side and the non-linearity is excitatory, leading to an enhancement of signals moving in the preferred direction (PD). In the full HR detector (Figure 1Aiii), two of those subunits, or half-detectors, are arranged in a mirror-symmetric fashion and subtracted from each other to yield a fully opponent detector (for review, see [8]).

How do the proposed elements of these algorithmic models map onto the neural circuits of the fly, and how does direction selectivity arise? The fly optic lobe consists of four neuropils downstream of the retina: the lamina, medulla, lobula, and lobula plate (Figure 1B). Photoreceptors synapse onto lamina monopolar cells. These lamina cells feed into two separate pathways encoding for different contrast polarities [9–11]: the ON pathway encodes brightness increments, and the OFF pathway encodes brightness decrements. In each pathway, the direction of visual motion is computed separately [12, 13]. In both pathways, lamina neurons connect onto a distinct set of medulla neurons. In the ON pathway, these medulla neurons have axon terminals in layer 10 of the medulla, where they overlap with the dendrites of T4 neurons [4]. In the OFF pathway, transmedulla neurons project to the lobula, where they synapse onto the dendrites of T5 neurons [5]. T4 and T5 neurons each fall into four subclasses, which respond selectively to visual motion in one of the four cardinal directions (front-to-back, back-to-front, up, and down) and project their axons according to this preference to one of the four layers of the lobula plate [14]. There, T4 and T5 cells converge and provide direct excitatory cholinergic input onto wide-field lobula plate tangential cells [15]. In addition, T4 and T5 cells synapse onto lobula plate intrinsic (LPi) neurons, which in turn inhibit tangential cells in the adjacent, oppositely tuned layer [16], making tangential cells fully motion opponent. Hence, T4 and T5



(legend on next page)

neurons would represent the half-detector units of the fully opponent motion detector model just before the subtraction stage. Although the HR detector describes the responses of lobula plate tangential cells well, the responses of T4 and T5 neurons are more directionally selective than would be expected for the half-detectors of the HR model [14, 17].

In the ON pathway, medulla intrinsic neuron 1 (Mi1) and transmedullary neuron 3 (Tm3) were originally suggested as the main inputs onto T4 neurons from electron-microscopic reconstructions [4]. These data showed a small spatial offset of about a fifth of a column, about 1° in visual space, between Mi1 and Tm3 synapsing onto the same T4 neuron, with Tm3 located toward the null side of the T4 neuron. Based on this spatial offset, two possible implementations of the motion detector were suggested: a HR correlator with Tm3, or a BL detector with Mi1, as the delayed arm. Subsequent patch-clamp recordings showed a small temporal delay of ~20 ms for Mi1 with regard to Tm3, as well as a similar temporal offset of Tm1 with respect to Tm2 in the OFF pathway [18]. This led to the suggestion of HR correlator implementations with Mi1 and Tm1 as the delayed and Tm3 and Tm2 as the direct arms in the ON and the OFF pathway, respectively [18, 19].

However, new findings from several recent studies question this model. First, new electron-microscopic circuit reconstructions show additional synaptic input from Mi4 and Mi9 cells onto T4 cells (Lou Scheffer, personal communication; https://web.archive.org/web/20150218101857/http://emanalysis.janelia.org/flyem_tables.php), and from the transmedulla neurons Tm4 and Tm9 onto T5 cells [5]. Second, when all four input cell types in the OFF pathway were considered, large differences in their temporal response kinetics to flashes of dark bars were revealed [20]. Whereas Tm1, Tm2, and Tm4 respond like band-pass filters with different time constants, Tm9 has the response characteristic of a pure low-pass filter, together forming a filter bank that lends itself well to the construction of motion detectors. Third, whereas blocking the synaptic output of Mi1 severely reduces responses of tangential cells to moving ON edges, blocking Tm3 output only affects responses to edges moving at higher angular velocities but leaves responses to lower velocities unchanged [21]. This again argues against Tm3 being one of the two arms of the motion detector under all conditions. Similarly, in the OFF pathway, all four cell types were shown to contribute to the detection of moving OFF edges. Blocking their synaptic output decreased the responses of downstream tangential cells and reduced the optomotor response to OFF edges [20]. However, no blocks of single cell types or of two types in combination fully abolished the responses to dark edges, suggesting either redundancy or a more complicated implementation than previously suggested. Fourth, recent experiments based on the sequential stimulation of individual laminar cartridges revealed that the elementary motion detectors in the ON pathway, T4 neurons, implement ND suppression [17] in addition to PD enhancement [22] (Figure 1Aiv). Spatiotemporal receptive fields of T5 neurons are consistent with a similar model in the OFF pathway [23]. This more elaborate motion detector implementation could explain the high direction selectivity. However, in contrast to both HR and BL detectors, it relies on at least three input elements.

Taken together, in both pathways, evidence mounts for a neural implementation that is more complicated than either the BL or the HR model alone, and there is a multitude of combinations possible to place the known columnar input elements into the proposed algorithmic three-arm model of the *Drosophila* motion detectors.

In order to dissect the roles and contributions of individual cell types, it would be helpful to modify their temporal response dynamics and observe the effect on the downstream motion detectors. One remarkable property of tangential cells is that their velocity tuning is not fixed but dependent on the behavioral state of the fly, as has been observed in *Drosophila* and *Lucilia*. In walking [24] as well as in tethered flying flies [25, 26], the temporal-frequency tuning shifts toward higher frequencies, corresponding to higher velocities, potentially matching the expected change of the stimulus statistics from resting to active locomotion. The behavioral effect can be mimicked in resting flies by pharmacological activation of octopamine receptors with octopamine [26] or the octopamine agonist chlordimeform (CDM) [25, 27]. The physiological source of this neuromodulation is octopaminergic neurons that project to the medulla, lobula, and lobula plate [28, 29]. They become activated during flight and are both necessary and sufficient for the increase in responses to higher temporal frequencies [26]. Importantly, this change in the temporal tuning could be reproduced in computer simulations by decreasing the low-pass filter time constant in the HR detector [25], indicating that identifying the input elements that change their kinetics under octopamine activation might help to pinpoint their functional roles in the detector.

Figure 1. Theoretical Models for Visual Motion Detection and the Underlying Neuronal Circuitry

(A) Algorithmic models of motion detectors based on variations of a common theme of spatiotemporal correlations of local luminance changes detected by photoreceptors. (Ai) In the Hassenstein-Reichardt (HR) correlator (of which a half-detector is shown here), a delay (τ) on the first of two arms activated by motion in the preferred direction (PD) causes coincidence of the two signals from neighboring photoreceptors (separated by an angle, $\Delta\phi$). A multiplicative non-linearity results in a PD enhancement. (Aii) In the Barlow-Levick (BL) detector the delay is located on the opposite arm, and the non-linearity is suppressive/inhibitory, causing a null-direction (ND) suppression. (Aiii) In the full HR correlator, two mirror symmetric subunits from (Ai) are subtracted, resulting in a fully opponent detector, which not only depolarizes in PD but also hyperpolarizes in ND. (Aiv) A recently proposed model, based on the responses of T4 neurons to apparent motion stimuli, combines PD enhancement and ND suppression along the PD axis.

(B) Schematic of the circuitry of the *Drosophila* optic lobe showing neuron classes suggested to be involved in visual motion detection. Local luminance changes are detected by photoreceptors in the retina and relayed via lamina monopolar neurons (classes L1–L5) and medulla neurons (Mi1, Tm3, Mi4, Mi9, Tm1, Tm2, Tm4, and Tm9) to T4 and T5 neurons. The latter are the first neurons in the visual pathway that respond selectively to motion. Both T4 and T5 form four subtypes that respond to one of the cardinal directions and project accordingly to the four layers of the lobula plate, thus forming a map of visual motion directions. In the lobula plate, they synapse onto large-field tangential cells (horizontal system [HS] and vertical system [VS] cells), as well as onto lobula plate intrinsic (LPi) cells that in turn form inhibitory synapses onto tangential cells in the adjacent layer of opposite PD. This inhibition corresponds to the subtraction stage in the full HR correlator (Aiii) and endows lobula plate tangential cells with full motion opponency.

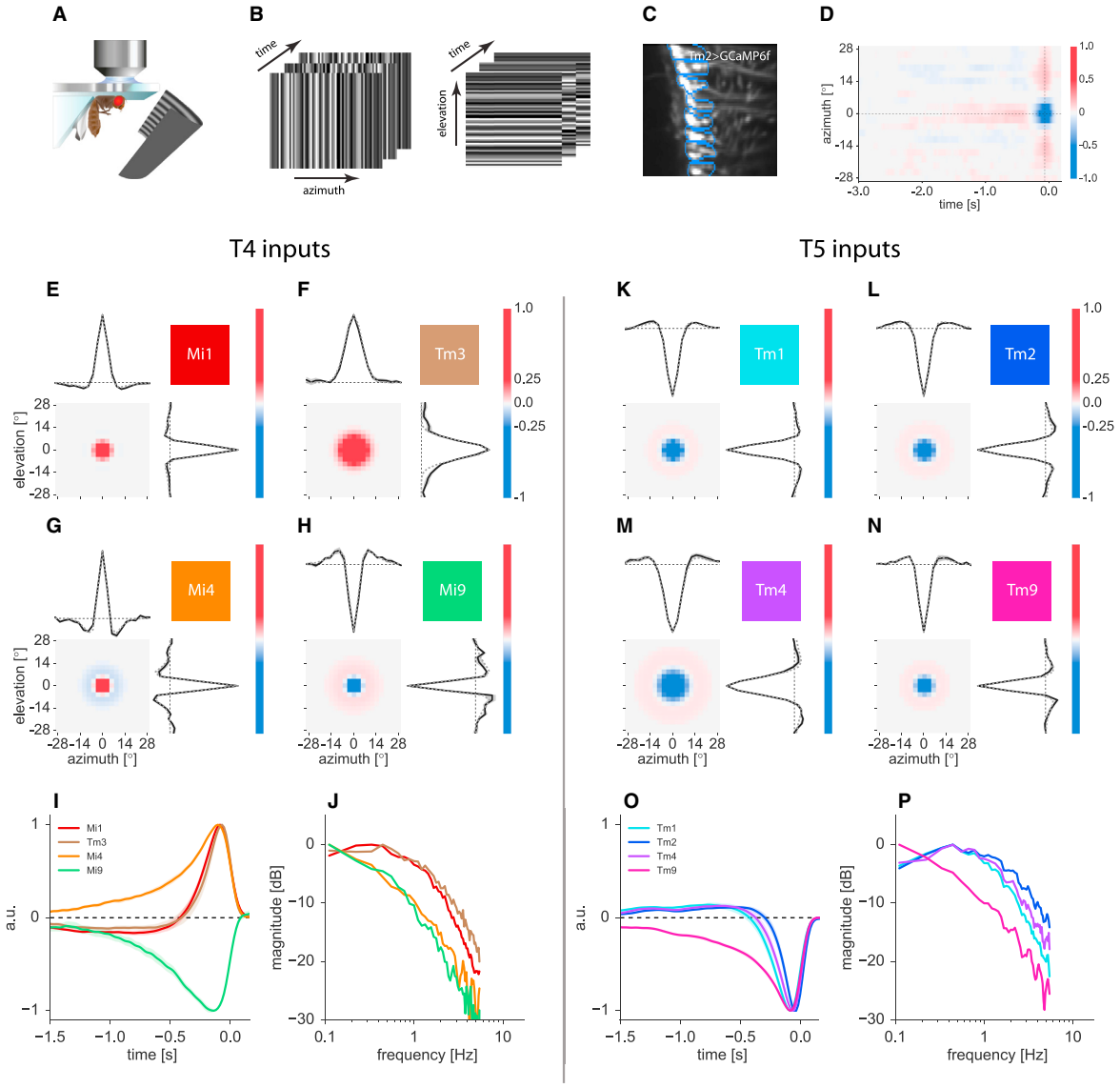


Figure 2. Response Properties of the ON- and OFF-Pathway Medulla Columnar Elements

(A) Two-photon calcium imaging of immobilized flies.

(B) Schematic of vertical (left) and horizontal (right) white-noise stimulus illustrated by three frames.

(C) Terminals of Tm2 neurons expressing the genetically encoded calcium indicator GCaMP6f. Regions of interest (ROIs) for the analysis of calcium indicator fluorescence changes encompass single terminals.

(D) Average aligned spatiotemporal receptive field of all Tm2 cells from (C) for a white-noise stimulus consisting of vertical bars. Along the vertical axis, the center-surround structure of the OFF-center receptive field is visible in the heat color code (vertical dashed line at the time of the peak of the response). The section along the time axis through the receptive field center reveals the temporal response kernel.

(E–H) Receptive fields of Mi1 (E), Tm3 (F), Mi4 (G), and Mi9 (H) for vertical (upper left) and horizontal (lower right) white-noise bar stimulation. From these, the two-dimensional receptive fields were constructed as a two-dimensional difference of Gaussians (Supplemental Experimental Procedures).

(I) Temporal kernels resulting from the reverse correlation of the calcium response with the white-noise stimulus for Mi1, Tm3, Mi4, and Mi9.

(J) Temporal kernels in frequency-space (constructed from the temporal kernels in (I) revealing Mi1 and Tm3 as band-pass filters and Mi4 and Mi9 as low-pass filters. (For the measurements of the spatial receptive fields: Mi1: N = 5 flies, n = 35 cells; Tm3: N = 6, n = 37; Mi4: N = 5, n = 33; Mi9: N = 7, n = 29. For the determination of the temporal kernels twice as many measurements, from the horizontal and vertical one-dimensional noise stimulus, could be used.)

(legend continued on next page)

In this study, we comprehensively characterize the spatiotemporal response profiles of all known columnar input elements of both the ON and OFF motion detectors in the fruit fly *Drosophila melanogaster* and take advantage of the motion detectors' state-dependent tuning characteristics. Using computer simulations, we test which combinations of input elements result in the observed properties of T4 and T5 neurons and thereby narrow down their possible cellular implementation. In particular, we address the question of whether the response dynamics of the input elements are sufficient to yield realistic motion detectors, or whether additional mechanisms on the synaptic or dendritic level are required to further modify the dynamics of the input signals.

RESULTS

Characterization of the Columnar Input Neurons to T4 Cells

The functional role of the input neurons to the elementary motion detectors and their correspondence to elements of any detector model depend crucially on their spatiotemporal response characteristics. For this reason, we characterized the spatial extent of the receptive fields as well as the response dynamics of all putative input elements to the T4 and T5 cells: Mi1, Tm3, Mi4, and Mi9 in the ON pathway, and Tm1, Tm2, Tm4, and Tm9 in the OFF pathway. Expressing the genetically encoded calcium indicator GCaMP6f [30] with cell-type-specific Gal4-driver lines, we imaged calcium signals in single terminals in layer 10 of the medulla or the proximal lobula for the ON- and OFF-pathway elements, respectively.

To precisely map the receptive fields of the input elements, we used a one-dimensional white-noise stimulus consisting of 2.8° wide horizontal or vertical bars covering the full extent of the arena (Figures 2A–2D; Figure S1; Supplemental Experimental Procedures). The spatiotemporal receptive fields were then determined from the neuron's calcium response by reverse correlation. The spatial components of these are the one-dimensional horizontal and vertical projections of the underlying two-dimensional spatial receptive field of the cell. In all cases, they strongly resembled a difference of Gaussians (DOG; also called a "Mexican hat"). Because they were similar for both the horizontal and vertical dimensions, we fitted a two-dimensional DOG to reconstruct two-dimensional spatial receptive fields (Figures 2E–2H and 2K–2N). The temporal component of the spatiotemporal receptive field reflects the temporal filtering properties of the neuron (impulse response). The extracted temporal filters were validated by predicting held-out test sequences of neuronal responses from the stimulus for two example neuron types (Mi1 and Tm9) (Figure S2; see Supplemental Experimental Procedures).

All four cell types in the ON pathway, Mi1, Tm3, Mi4, and Mi9, showed locally confined receptive fields that appeared isotropic in the horizontal and vertical dimensions (Figures 2E–2H). Mi1, Mi4, and Mi9 cells revealed a receptive field center with a half-width diameter of approximately 6°–7°, corresponding to about one optical column. In contrast, the receptive field center of Tm3 was substantially larger, with a half-width diameter of about 12°. Mi4 and Mi9, and to a lesser degree Mi1, also revealed a significant antagonistic surround, giving them spatial band-pass characteristics. This antagonistic surround had a half-width diameter of approximately 20° for both Mi4 and Mi9 (Table S1). Because the area and thus the volume under the curve are proportional to the square of the radius, the amplitude ratio of surround to center should equal the inverse of the ratio of the squares of their half-widths for the center and the antagonistic surround to cancel perfectly. Notably, this relation is fulfilled for both low-pass elements, and the integrals of their surrounds perfectly match their respective centers, thus predicting no responses to wide-field flicker stimuli. At the same time, the spatial band-pass filter enhances responses to edges within the visual scene. In the case of Mi1, the integral of the surround reached about 50% of the one of the center. For Tm3, surround inhibition was completely absent, such that those cells have a pure low-pass characteristic in the spatial domain.

The temporal component of the spatiotemporal receptive field centers yielded the impulse responses, which reflect the temporal filtering properties of the respective cell type. Mi1 and Tm3 showed band-pass filter characteristics, as can be seen in their biphasic impulse responses (Figure 2I) and in their response spectra (Figure 2J). In contrast, Mi4 and Mi9 appeared as pure low-pass filters (Figures 2I and 2J). Surprisingly, and in contrast to the other elements of the ON pathway, Mi9 showed the inverse contrast preference, with an increased calcium response to darkening in its receptive field center (OFF response). However, apart from the polarity, the time course and filter characteristics of Mi9 were very similar to those of Mi4 (Figures 2I and 2J). Thus, the four ON-pathway elements can essentially be grouped into two classes: two fast-transient cells (Mi1 and Tm3) and two slow-sustained cells (Mi4 and Mi9). Within each class, the cells' impulse responses revealed only small differences.

Characterization of the Columnar Input Neurons to T5 Cells

We next performed analogous experiments on the OFF-pathway elements Tm1, Tm2, Tm4, and Tm9. Mirroring the situation in the ON pathway, all four neurons of the OFF pathway had locally confined isotropic receptive fields (Figures 2K–2N). In agreement with previous reports [18, 20], they were all excited by luminance decrements. Accordingly, they revealed an OFF receptive field center. The receptive fields of all four cells also had an antagonistic surround component, giving them a spatial band-pass characteristic. In contrast to Mi4 and Mi9, however, the surround inhibition, with respect to the center, was weaker, which should render them more responsive to wide-field flicker. As a parallel to the ON-pathway elements, three of the neurons, Tm1, Tm2, and Tm9, showed a receptive field center with a half-width diameter of about 7°, whereas one element, Tm4, had a larger receptive field center, with a half-width diameter of approximately 10°.

(K–P) Characterization of the inputs to T5 cells in the OFF pathway. Spatial receptive fields of Tm1 (K), Tm2 (L), Tm4 (M), and Tm9 (N). Temporal kernels in the time (O) and frequency domain (P) for the four input elements in the OFF pathway. (Tm1: N = 8 flies, n = 71 cells; Tm2: N = 9, n = 93; Tm4: N = 5, n = 35; Tm9: N = 5, n = 32.)

Graphs depict the mean. Shaded areas around the line, where displayed, represent ±SEM. See also Figures S1 and S2 and Tables S1 and S2.

The half-width of the antagonistic surround amounted to about 25° for Tm1, Tm2, and Tm9 and to 35° for Tm4 (Table S2). As for Mi1, and in contrast to Mi4 and Mi9, the antagonistic surround strength for all OFF input elements reached about 50% of the center, as calculated above on the basis of the amplitude and half-width ratios.

As for the ON-pathway elements, we assessed the temporal filter dynamics by measuring the impulse responses within the receptive field centers (Figures 2O and 2P). This revealed a clear band-pass characteristic for Tm1, Tm2, and Tm4 with rather short low-pass time constants of about 100–270 ms. In contrast, the impulse response of Tm9 reflected a pure low-pass filter with a much longer time constant of about 500 ms. Within the group of band-pass filters, Tm1, Tm2, and Tm4 responses have different time courses (Figure 2O) and response spectra (Figure 2P), corroborating a previous study [20]. Thus, as a striking difference from the ON-pathway elements, where two fast and two slow cells are found, the OFF pathway comprises three fast and only one slow cell.

Application of the Octopamine Agonist CDM Changes the Temporal Frequency Tuning of T4 and T5 Cells

It has previously been shown that activation of the octopamine system modulates the temporal-frequency tuning of lobula plate tangential cells [25, 26]. This effect could be implemented directly at the level of the tangential cells, or indirectly, by modifying the temporal tuning properties of its presynaptic input neurons, i.e., the T4/T5 cells. The latter case would give a handle to manipulate the elementary motion detectors and potentially allow narrowing down of their cellular implementation.

We first confirmed that the activation of the octopamine system with the octopamine agonist CDM [31] at a concentration of 20 μM [25] shifts the temporal tuning of tangential cells in the lobula plate of immobilized *Drosophila* to higher frequencies (Figure S3), corroborating earlier findings using octopamine [26].

Next we focused on T4 and T5 neurons. We performed two-photon Ca^{2+} imaging in *Drosophila* expressing the genetically encoded calcium indicator GCaMP6m in the subset of T4/T5 neurons that are upward motion selective and project their axons to layer 3 of the lobula plate (T4c/T5c) (Figure 3A). Visual stimulation was presented on a semi-cylindrical LED arena and consisted of full-contrast square-wave gratings with a spatial wavelength of 24°, moving at 12 different velocities ranging from 1.2°/s to 480°/s, corresponding to temporal frequencies from 0.05 to 20 Hz, in PD and ND. Responses of T4 and T5 neurons were quantified as relative change of fluorescence (DF/F) amplitudes within small regions of interest in lobula plate layer 3 (example traces in Figure 3B). We found a temporal-frequency optimum of 1 Hz for motion in PD (Figure 3C, black traces). Application of CDM shifted the temporal-frequency optimum from 1 Hz in control to about 2.5 Hz (Figure 3C, magenta traces). Recording Ca^{2+} signals from the dendrites of either T4 or T5 cells, we found that T4 and T5 cells, considered separately, exhibited a similar temporal-frequency tuning, under control conditions as well as after application of CDM, and a similar shift in their tuning with CDM (Figures 3D and 3E).

In order to distinguish changes in the response to isolated motion stimuli from changes in the temporal integration of periodic signals, we also tested the velocity tuning of T4 and T5 neurons to moving edges. For this, we presented bright and dark edges of full contrast moving at different speeds ranging from 3°/s to 300°/s in PD (Figures 3F and 3G). Corroborating previous results [14], T4 neurons responded selectively to bright edges, whereas T5 neurons were found to be selective for motion of dark edges. Measuring the calcium responses in the axon terminals in the lobula plate, we found that under control conditions the responses were highest to edges moving at the slowest velocity of 3°/s for both ON and OFF edges, i.e., T4 and T5 neurons, respectively (Figures 3F and 3G, black traces). As was seen for the grating stimuli above, application of CDM shifted the optimal stimulus condition to higher velocities of 12°/s (Figures 3F and 3G, magenta traces).

Therefore, the shift of the temporal tuning properties of lobula plate tangential cells during flight or mimicked by the application of octopamine-receptor agonists (Figure S3 [25, 26]) is already present at the level of the T4 and T5 cells, thus affecting the tuning of the elementary motion detectors.

Octopamine-Receptor Activation Speeds the Input Elements of T4 and T5 Cells

Different possible mechanisms could explain this shift of temporal tuning in T4/T5 cells. On the one hand, octopamine signaling could affect the synaptic inputs onto T4 and T5 neurons by changing the kinetics of neurotransmitter receptors or the dendritic integration of those signals in T4/T5 neurons. Different input elements with different response kinetics could differentially contribute to the postsynaptic signals in different states through changes in their response amplitude or via their synaptic weight. On the other hand, the kinetics of some or all input elements could speed up. We set out to test the latter hypothesis, i.e., that the response characteristics and tuning of the elementary motion detectors result directly from the temporal dynamics of the respective input elements.

For this, we characterized the spatiotemporal receptive fields of all input elements in both the ON and OFF pathways after activation of the octopamine system with CDM and compared them to control conditions. Application of CDM left the spatial receptive fields of all four input neurons in the ON pathway unaffected (Figure 4A). However, it accelerated the response kinetics of all four cell types to different degrees, with much stronger effects on the fast band-pass elements Mi1 and Tm3 than on the slow low-pass filters Mi4 and Mi9 (Figures 4B and 4C, magenta traces; Figures S4A, S5Ai, and S5Bi). As for control conditions, response kinetics of Mi1 and Tm3, as well as of Mi4 and Mi9, remained similar to each other after addition of CDM. In the OFF pathway, the results were very similar. The spatial receptive fields appeared unchanged by CDM for any of the columnar input neurons (Figure 4D). However, in the temporal domain, addition of CDM to the bath sped up the impulse responses significantly (Figures 4E and 4F, magenta traces; Figures S4B, S6Ai, and S6Bi), as was seen in the ON-pathway band-pass elements.

Computer Simulations Based on the Input Elements' Temporal Filters Suggest Candidate Motion Detectors

The input elements to the motion-detecting neurons T4 and T5 can be roughly grouped into two classes: temporal low-pass filters with large time constants, and band-pass filters with

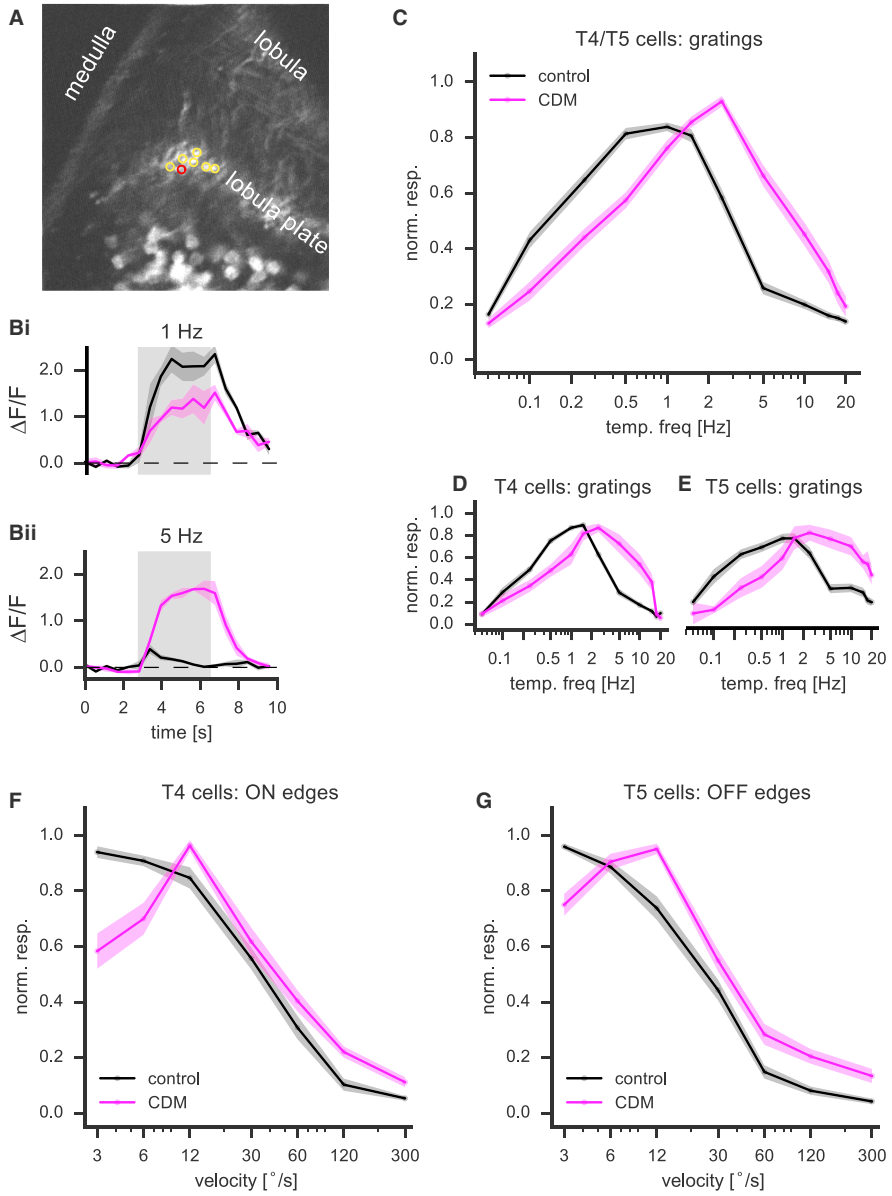


Figure 3. Application of CDM Shifts the Temporal-Frequency and Velocity Tunings of T4/T5 Cells to Higher Velocities

(A) T4/T5 neurons of the upward motion-selective subtype "c" projecting their axons to layer 3 of the lobula plate, expressing the genetically encoded calcium indicator GCaMP6m. The circles mark ROIs in the lobula plate; the red circle corresponds to the example calcium traces in (B).

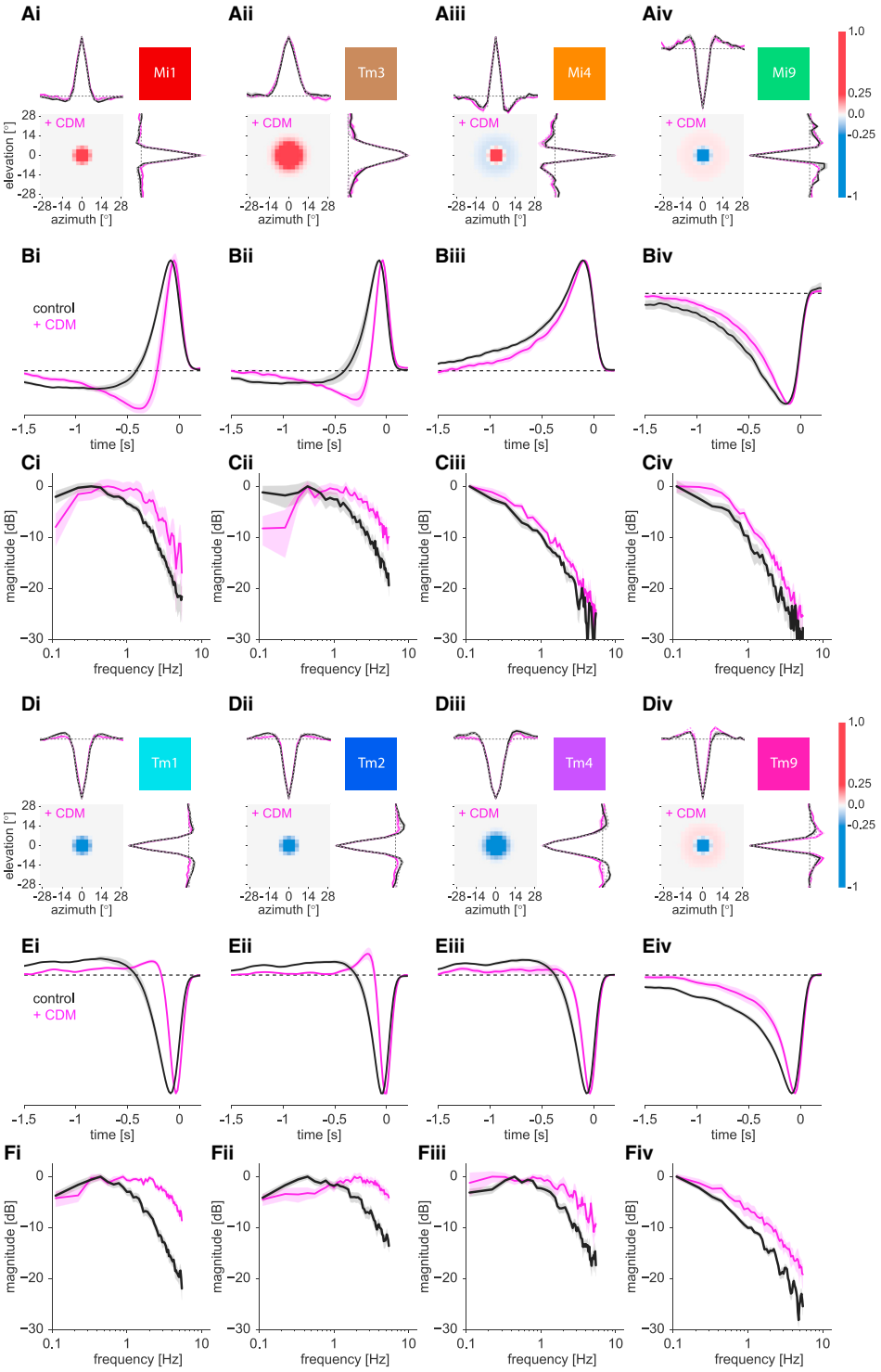
(B) Example of calcium responses (fluorescence changes) in the axon terminals of T4/T5 cells in response to square-wave gratings moving at temporal frequencies of 1 Hz (Bi) and 5 Hz (Bii) in control (black) and after application of CDM (magenta).

(C) Population average of responses of T4/T5 axon terminals to square-wave gratings moving in the PD (up). Application of CDM leads to a shift of the temporal tuning optimum (N_{cont} = 36 flies, n_{cont} = 80 ROIs; N/n_{CDM} = 15/39).

(D and E) Characterization of the temporal-frequency tuning in T4 (D) and T5 dendrites (E). As observed for the axon terminals, application of CDM (magenta) shifts the temporal-frequency tunings of both T4 and T5 cells to higher frequencies, as compared to control (black) (T4: N_{cont} = 27 flies, n_{cont} = 52 ROIs, N/n_{CDM} = 9/14; T5: N/n_{cont} = 18/27, N/n_{CDM} = 7/9).

(F and G) Population average of responses of T4 and T5 axon terminals in the lobula plate to bright (F; T4) and dark edges (G; T5), moving at different velocities, in control (black) and after application of CDM (magenta) (N_{cont} = 9 flies, n_{cont_T4} = 21, n_{cont_T5} = 37 ROIs; N_{CDM} = 6, n_{CDM_T4} = 16, n_{CDM_T5} = 17).

Graphs depict the mean. Shaded areas around the line represent ±SEM. See also Figure S3.



(legend on next page)

significantly shorter time constants. We used the above-determined spatial receptive fields and response kinetics of the input elements and asked whether these could predict the responses of their postsynaptic targets, the elementary motion detector T4/T5 cells, without the necessity of additional filters or delays implemented either at the level of the synapses between the inputs and the T4/T5 cells or within the dendrites of the T4/T5 cells itself. In addition, we asked whether the observed shift in the temporal tuning in T4/T5 cells after application of the octopamine agonist CDM could be fully explained by the change of filter properties of the respective input neurons.

Although GCaMP6f has relatively fast kinetics when compared with other calcium indicators, it still possesses a decay time constant on the order of hundreds of milliseconds [30, 32]—long enough to significantly prolong the calcium signals of cells that have temporal dynamics on the same order of magnitude. In order to correct for this temporal filtering by the calcium indicator itself, we deconvolved the impulse responses in the frequency domain with a GCaMP6f low-pass filter (Figures S5 and S6). These corrected spectra were used as an approximation of the underlying filter properties of the input cells by fitting first-order filters to the average corrected frequency responses (Tables S1 and S2). We then used these values as well as the spatial filter characteristics in our computer simulations of a motion detector. Because the synaptic transmitters and postsynaptic receptors, and therefore the sign of the synaptic inputs, are not known, we decided not to make any assumptions about the sign of the synapses and ignored the response polarities of the determined receptive fields in our simulations.

Our simulations were based on a motion detector that combines PD enhancement and ND suppression, resembling a hybrid of a HR half-detector and a BL detector, as described in Haag et al. [17] (Figure 1Aiv). In this detector, three inputs with receptive fields offset by 5° each along the PD axis are processed such that an enhancing input A on the null side (left) forms a multiplicative non-linearity with the central, direct input (B), whereas a suppressing input (C) on the preferred side (right) implements a divisive non-linearity. The response of this detector equals the product of the input signals on the enhancing and the direct arm, divided by the signal from the suppressing arm (see the Supplemental Experimental Procedures).

There are 24 possible permutations that map the four input elements of each pathway onto the three positions of this detector, each one resulting in a detector with different tuning properties. Without making any further assumptions, we asked whether some of these combinations would yield more direction-selective motion detectors than others. Each simulated detector was tested with moving square-wave gratings, and the responses were quantified in three ways (Figure 5A): (1) To assess how well the particular detector model discriminates between motion along PD and ND across velocities, we simulated square-wave gratings moving in PD and ND at different speeds covering more than three orders of magnitude. From the simulated responses, we calculated a direction selectivity index (DSI) as the relative difference between PD and ND responses, averaged over all grating velocities/temporal frequencies. (2) To judge the frequency tuning, we determined the temporal frequency evoking the maximum response in PD (temporal-frequency optimum, f_{opt}). (3) To characterize the direction tuning beyond PD and ND, emphasizing tuning sharpness, we simulated gratings moving in 12 equally spaced directions at the f_{opt} of each detector, as determined above. From those simulated responses, the normalized length of the tuning vector (L_{dir}) was calculated [33]. This tuning vector length of the hybrid detector was furthermore compared with the ones of the constituting HR and BL modules (Figures 1Ai and 1Aii, respectively).

In general, detectors with the low-pass filters Mi4 and Mi9 on both the outer enhancing and suppressing arms, flanking one of the band-pass elements Mi1 or Tm3, performed extremely well: they showed a rather high degree of direction selectivity and tuning sharpness (Figure 5B), in good agreement with the experimental data from T4 cells (compare with [14]), and their temporal-frequency optimum matched that of T4 cells as well (Figure 5B, right; compare with Figures 3C and 3D).

In addition, most combinations with one central low-pass neuron, Mi4, or, particularly, Mi9, flanked by the two band-pass elements Mi1 and Tm3, also achieved high direction-selectivity values. The PD (see arrows in Figure 5B, left) of these detectors is inverted as a consequence of the position of the delay in the HR and BL sub-modules. However, when considering both sub-modules separately (blue and red bars, respectively, in Figure 5B, right), the BL alone showed very low tuning sharpness (L_{dir}) and thus contributed little to the hybrid detector. This affects the tuning specificity of the hybrid detector, as can be seen when comparing, for example, Tm3xMi9/Mi1 with Mi9xTm3/Mi4. Both detectors are built on the same HR detector (using the same cells), but the one that employs Mi4 for the BL part of the model has a higher tuning sharpness. The same is true for all other pairs of this kind: given one pair of cells for the HR module, the implementation that places two low-pass filters on the outer arms of the detector always has the sharper tuning.

Figure 4. Activation of Octopamine Receptors Accelerates the Temporal Filters of the ON- and OFF-Pathway Medulla Columnar Elements

(A) Spatial receptive fields of Mi1 (Ai), Tm3 (Aii), Mi4 (Aiii), and Mi9 (Aiv) for vertical (upper left) and horizontal (lower right) white-noise bar stimulation under control conditions (black traces) and after application of CDM (magenta traces and two-dimensional receptive fields).

(B) Temporal kernels for Mi1 (Bi), Tm3 (Bii), Mi4 (Biii), and Mi9 (Biv) revealing the faster time course after application of CDM (magenta) as compared to control (black).

(C) Temporal kernels in frequency-space, constructed from the temporal kernels in (B). Application of CDM (magenta) leads to a shift of the center frequency of the band-pass filters as compared to control (black). (For the measurements of the spatial receptive fields [controls are as in Figure 2]: Mi1: $N_{CDM} = 5$, $n_{CDM} = 31$; Tm3: $N/n_{CDM} = 6/34$; Mi4: $N/n_{CDM} = 5/38$; Mi9: $N/n_{CDM} = 7/37$. Again, the temporal kernel results determined from the horizontal and vertical one-dimensional noise stimuli were pooled, resulting in twice as many measurements.)

(D–F) Analogous to (A)–(C), the spatial receptive fields (Di–Div), temporal kernels (Ei–Eiv), and frequency spectra (Fi–Fiv) of the OFF-pathway elements Tm1, Tm2, Tm4, and Tm9. (Controls are as in Figure 2; Tm1: CDM: $N_{CDM} = 8$, $n_{CDM} = 67$; Tm2: $N/n_{CDM} = 9/93$; Tm4: $N/n_{CDM} = 5/28$; Tm9: $N/n_{CDM} = 5/42$.)

Graphs depict the mean. Shaded areas around the line represent ±SEM. See also Figure S4 and Tables S1 and S2.

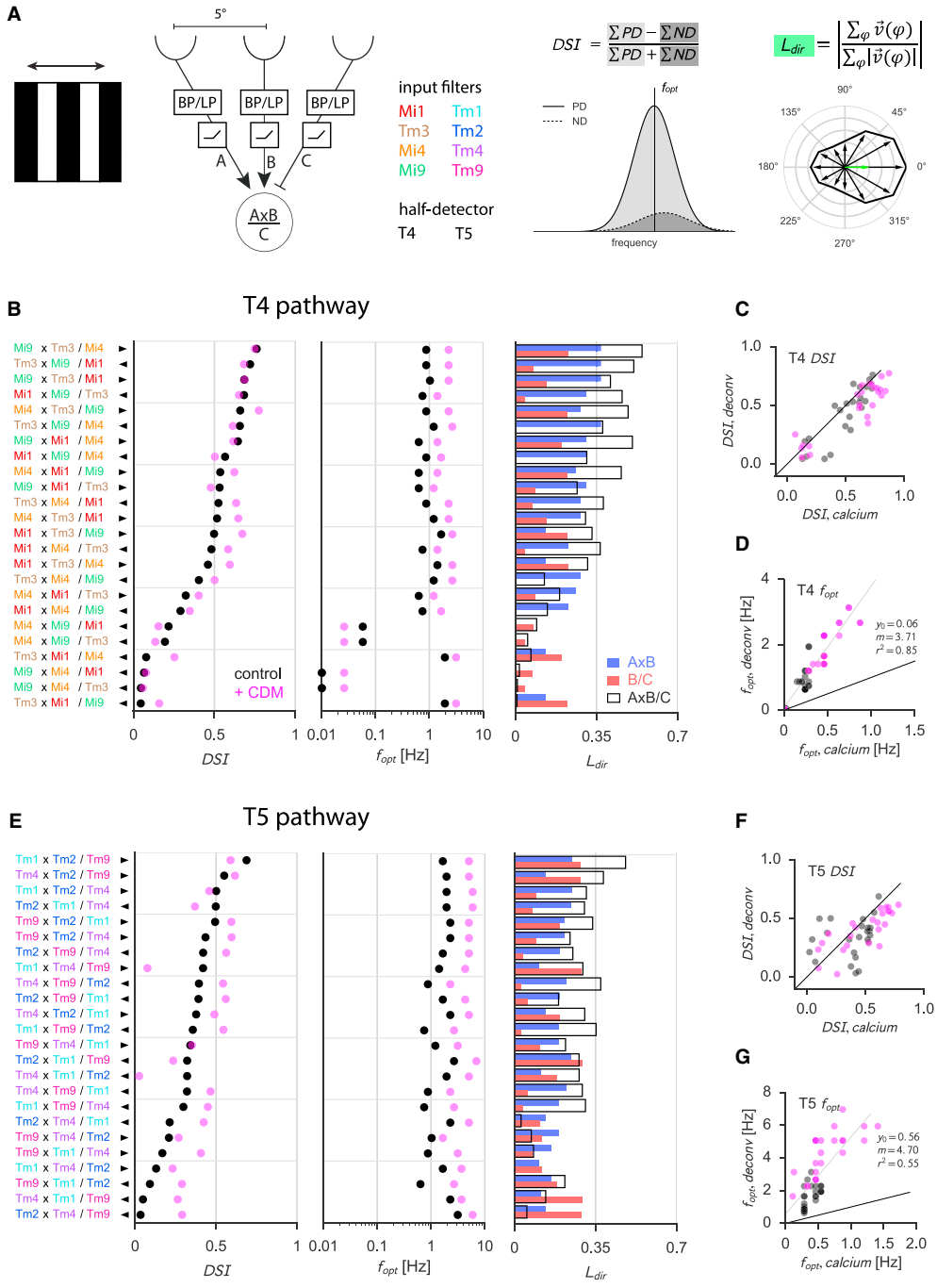


Figure 5. Computer Simulations of Elementary Motion Detectors

(A) Left: schematic of a three-arm detector combining a multiplicative PD enhancement and a divisive ND suppression. The positions of the enhancing ("A"), central ("B"), and suppressing ("C") input can be occupied by any but different input elements. Those input elements are described by their temporal filtering characteristics, implemented as a band-pass (BP) and low-pass filter (LP) with subsequent rectification. The receptive fields of the three inputs are offset by 5° each. The simulated detectors are stimulated with square-wave gratings moving at different temporal frequencies in PD and ND. Middle: the direction selectivity

(legend continued on next page)

Similarly, detectors that incorporated two elements with similar temporal response properties (such as Mi1 and Tm3) on two adjacent positions tended to perform worse, especially with respect to the L_{dir} value, indicating poor tuning sharpness. This can be easily explained by the fact that both the HR and BL modules of the hybrid detector rely on temporal differences in their respective two input arms. Inputs with more similar kinetics thus render the corresponding module less effective in creating direction selectivity. In fact, the best detectors were those where both halves showed high direction selectivities on their own (Figure 5B, right), provided the PDs of both modules were aligned.

Interestingly, almost all combinations showed a shift in their tuning toward higher temporal-frequency optima by about a factor of 2 when the filter properties after application of the octopamine agonist CDM were used, matching the experimental data. As a control that the direction selectivities in our simulations were not dependent on the used deconvolution filter, we repeated the simulations with the raw temporal kernels derived from the calcium responses. The same arrangements of input elements led to the motion detectors with the highest direction-selectivity values (Figure 5C), consistent with the notion that it is the relative filter properties that are crucial. Deconvolution merely changes the temporal frequency of the visual stimulus that leads to the maximum response (Figure 5D).

In the above simulations, we followed an unbiased approach with all inputs separated by 5°, thus having receptive fields arising from neighboring neuro-ommatidia. However, electron-microscopic reconstructions have shown a spatial offset between Tm3 and Mi1 cells projecting to the same T4 cell of about 1° in this order along the PD of the postsynaptic T4 cell [4]. The smaller spatial offset could counterbalance the small differences in temporal kinetics between these cells. Repeating the above simulations of the three-arm detector under these constraints still resulted in poorly direction-selective detectors for these combinations, with L_{dir} values of 0.38 (for Mi9xTm3/Mi1, as compared to 0.41 for a 5° offset) or less. In fact, when considering only a simple two-arm detector (HR or BL type), any detector that consisted of Tm3 and Mi1 with a spatial offset of 1° resulted in L_{dir} values of less than 0.06 for both types of detectors (in comparison to 0.13 for a 5° offset).

Although the evidence is weaker for the structure of the motion detector implementation in T5, we constructed analogous motion detectors for the OFF pathway with the measured receptive fields and response kinetics of the columnar inputs onto T5 neurons (Figure 5E). In contrast to the ON pathway, only one out of the four input elements, Tm9, constitutes a low-pass filter, whereas the other three, Tm1, Tm2, and Tm4, exhibit band-pass characteristics. Most input element combinations resulted in motion detectors with low direction selectivity. Notably, the highest direction selectivity resulted from detectors with the low-pass filter Tm9 on the suppressing arm. Naturally, detectors with the fastest input (principally Tm2) in the central position flanked by two slower elements achieved higher direction selectivities, as with this arrangement the PDs of the HR and BL subunits are aligned. Arrangements with the sole low-pass filter, Tm9, in the central position resulted in detectors with poor directional tuning, both measured as DSI across all frequencies and L_{dir}, resulting from a virtually ineffective BL half (Figure 5E, right). Interestingly, combinations with the band-pass filters Tm1 and Tm4 constituting either half of the detector tended to perform comparatively poorly—and sometimes even showed a complete breakdown of direction selectivity—in at least one of the simulated physiological states. This can be explained by the fact that the small differences in the temporal response kinetics of these neurons were not stable between control and under CDM (Figure S6). As was seen for the ON pathway, using the spatiotemporal filters extracted under CDM in the simulations led to an increase of the temporal-frequency optimum by about a factor of 2 across all detectors (Figure 5E, middle, magenta dots). Again, the simulations were robust to the deconvolution applied to account for the filtering by the calcium indicator (Figure 5F). The best arrangements were the same irrespective of whether the raw or deconvolved filters were used, and only the temporal-frequency optimum was affected (Figure 5G).

Taken together, we find distinctly different response kinetics of the input elements in both the ON and the OFF pathway, from band-pass filters to pure low-pass filters. These map naturally onto hybrid elementary motion detectors implementing PD enhancement and ND suppression. The best-performing detectors arise when the fastest element occupies the central arm, flanked by slower inputs on the enhancing and suppressing arms. In the ON pathway, two low-pass inputs, Mi4 and Mi9, are found to fill this role. In the OFF pathway, the single low-pass element, Tm9, appears to be best positioned on the suppressing arm to achieve the highest direction selectivity.

DISCUSSION

To understand how motion detection is implemented on the dendrites of T4 and T5 cells, we describe in this study the response

of the detector is assessed across all temporal frequencies based on the area under the temporal-frequency tuning curves in PD and ND as the direction selectivity index: DSI = (ΣPD − ΣND) / (ΣPD + ΣND). The dotted line indicates the temporal-frequency optimum (f_{opt}) for responses in PD. Right: illustration of the normalized tuning vector length (L_{dir}) as a measure for direction selectivity and tuning sharpness. L_{dir} is calculated as the vector sum of all responses according to the direction of stimulus motion, normalized to the sum of all response vector lengths.

(B–D) Characterization of the simulated motion detectors for the ON pathway.

(B) Direction selectivity (left), temporal-frequency optimum (middle), and normalized tuning vector length (right) for all possible permutations of the four ON-pathway input elements on the three positions of the simulated detector. The magenta dots indicate the effect of CDM application on direction selectivity and temporal-frequency tuning resulting from the accelerated temporal filters of the input elements. Arrows indicate the PD with respect to the corresponding cell arrangements. For the tuning vector length, the hybrid detectors (black open bars) were compared to their constituting HR ("AxB"; blue) and BL modules ("B/C"; red).

(C and D) Direction-selectivity indices (C) and temporal-frequency optima (D) of all detectors based on the deconvolved filter kernels as shown in (B) plotted against the detectors based on the raw calcium kernels.

(E–G) Same as (B)–(D) but for the OFF pathway.

See also Figures S5 and S6.

properties of the elementary motion detectors in *Drosophila*, the T4 and T5 neurons, as well as all of their known columnar synaptic input neurons, under two different tuning regimes. With this comprehensive characterization, we are able to narrow down the cellular implementation of the motion detectors and suggest probable wiring diagrams.

All of these input elements possess spatially restricted receptive fields with centers spanning one to two ommatidia. All, with the exception of the ON-pathway band-pass neuron Tm3, have pronounced antagonistic surrounds. Particularly for the low-pass filter elements Mi4 and Mi9, the strong antagonistic surround fully counterbalances the excitatory center. This should not only eliminate sensitivity to large-field flicker stimuli but more importantly curtail the otherwise tonic responses of pure low-pass filters to moving edges, and thus strongly improve direction selectivity. The locally confined receptive fields are in agreement with previous studies [18, 20, 34] but in contradiction to [35], which described Tm9 as a wide-field neuron. In both pathways, one neuron shows a larger receptive field (Tm3 in the ON pathway, and Tm4 in the OFF pathway). The larger receptive field sizes of Tm3 [18] and Tm4 neurons are consistent with the multi-columnar input these neurons receive based on electron-microscopic reconstructions [4, 5].

All elements of the OFF pathway respond to light OFF in the center of their receptive fields, consistent with [20]. In the ON pathway, Mi1, Tm3 [18], and Mi4 analogously show an ON-center response. Mi9, however, despite being an element in the ON pathway, responds positively to OFF stimuli. This could suggest a sign reversal through an inhibitory synapse onto T4. However, it is not known what neurotransmitter is released by Mi9, and thus whether it excites or inhibits T4 neurons.

Within each of the two pathways, we find a diversity of temporal filter characteristics from fast band-pass filters to pure low-pass filters with slow-sustained responses. These differences in temporal dynamics make them ideal components for motion detection without the need of postulating further processing by slow synaptic signaling or electrotonic filtering within the dendrites of T4 and T5 cells. Where the response kinetics of these cells have been previously described, our data are consistent. In particular, Mi1, Tm3, Tm1, Tm2, and Tm4 have previously been shown to respond transiently to sustained stimuli, i.e., to possess band-pass characteristics [18, 20, 34]. Tm3 appears faster than Mi1 [18] (but see [36]), and Tm2 faster than Tm1 [18, 20, 36]. However, these temporal differences are often very small. On the other hand, Tm9 in the OFF pathway has been described as a low-pass filter [20, 35], which matches our results. In the ON pathway, we find that the previously uncharacterized cell types Mi4 and Mi9 also show pure temporal low-pass response characteristics. Thus, in both pathways, input elements with slow-sustained and fast-transient responses are found, which then converge onto the dendrites of T4 and T5 cells, respectively. Yet the relative distribution differs. In the ON pathway, two input elements show pure low-pass characteristics (Mi4 and Mi9), whereas in the OFF pathway, Tm9 constitutes the only pure low-pass filter. Two of the three input elements that constitute pure low-pass filters, namely Mi9 in the ON and Tm9 in the OFF pathway, receive their lamina input primarily from the lamina monopolar neuron L3 [37]. As L3 has been shown to respond in a slower and more sustained fashion [38] than, e.g., the transient L2 [10, 11], this could explain the low-pass characteristics of Mi9 and Tm9. L3, like all lamina neurons, responds positively to light decrements, and it releases the excitatory neurotransmitter acetylcholine, explaining the OFF response of Tm9 and Mi9. The response dynamics of Mi4 are likely to be heavily shaped by the strong reciprocal connections with Mi9 [37]. These reciprocal connections, and thus likely the cells themselves, would have to be inhibitory, as these cells show opposite response polarities.

Based on the spatial receptive fields and response kinetics, we could ask which input neurons could play which role in the motion detector. Previous computer simulations based on the measured dynamics of Tm cells in the OFF pathway have shown that most combinations of two elements result in classical (full) HR detectors with similar temporal tuning optima roughly matching the tuning of tangential cells [20]. In that study, only the combination of Tm2 and Tm4 could be excluded, as their filter time constants were too similar to each other to result in a functioning detector. However, subtraction of oppositely tuned half-detectors not only leads to motion opponency but increases direction selectivity of otherwise poorly tuned half-detectors. Consequently, the tuning of lobula plate tangential cells represents a rather indirect readout. By comparing simulations of the half-detector stage with recordings from T4 and T5 neurons, we can exclude the majority of possible combinations of input elements based on their temporal-frequency optimum or directional selectivity (see below).

Based on visual stimulation of single individual columns, T4 neurons have recently been shown to implement both PD enhancement and ND suppression [17]. The receptive fields of these interactions are spatially offset along the PD axis in this order. The corresponding hybrid of an HR half-detector and a BL detector requires a minimum of three columnar inputs: a fast central input, flanked by two outer inputs providing signals that are delayed relative to the central one.

In our computer simulations for the ON pathway (Figure 5), the majority of detectors with the highest direction selectivity fall into two groups: (1) the two low-pass filter elements Mi4 and Mi9 on the outer enhancing and suppressing arms, and either of the fast band-pass elements Mi1 and Tm3 on the central arm, matching the above layout, and (2) the inverted arrangement, with one central low-pass filter, flanked by the band-pass filter elements Mi1 and Tm3. This also resulted in an inverted PD.

In the latter case, however, the BL subunit considered alone contributed very little to the directional tuning (Figure 5B, right), as the low-pass-filtered central excitatory input tends to outlast the corresponding suppression from the band-pass outer arm. This reduces the tuning sharpness of these detectors. Furthermore, this implementation does not match the arrangement of PD-enhancement and ND-suppression receptive fields along the PD in this order found for T4 cells [17]. Additionally, this arrangement would require Mi1 and Tm3 on the outer arms of the model, which is in stark contrast to their reported 1° spatial offset [4].

Among the more direction-selective detectors was also one combination with Tm3 on the central and Mi1 on the suppressing arm. However, the resulting BL subunit considered alone shows very poor directional tuning, and the direction selectivity arises

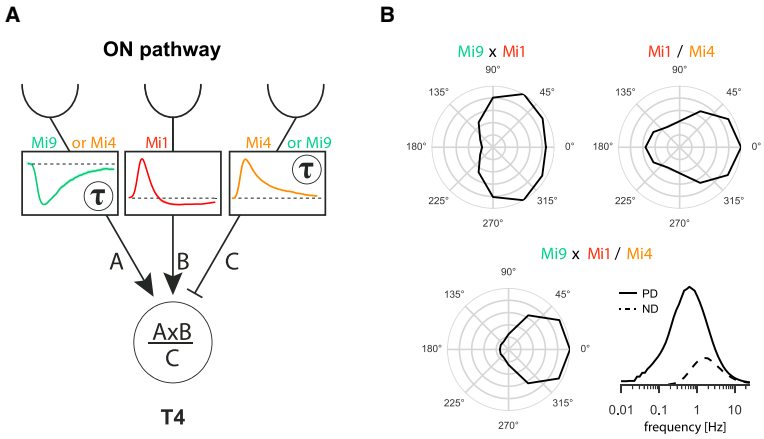


Figure 6. Proposed Implementation of the Elementary Motion Detectors in the ON Pathway

(A) T4 neurons implement both PD enhancement and ND suppression with receptive fields offset in this order along the PD axis. This requires one central fast arm being flanked by two delayed or stronger low-pass-filtered inputs. The relatively fast kinetics of Mi1 or Tm3 would suggest either or both for the central input. Mi4 and Mi9, on the other hand, show pure low-pass characteristics in their temporal kernels fitting the requirements of the two outer arms. The signs of both outer-arm synapses depend on the arrangement of Mi4 and Mi9 to accommodate their respective response polarity and match them to the required enhancing and suppressive inputs.

(B) Simulated detector responses for gratings moving across the visual field in 12 different directions, separated from each other by 30°. Top: directional tuning for the two sub-modules of this detector. Top left: the pure HR (half) detector Mi9xMi1 shows some direction selectivity but has a low tuning sharpness. Top right: the pure BL detector Mi1/Mi4 shows a substantial response in the ND direction (180°). Bottom: directional tuning for the hybrid detector Mi9xMi1/Mi4. This hybrid detector is very sharply tuned to rightward motion (left), whereas its direction selectivity remains high across stimulus frequencies (right).

virtually exclusively from the HR subunit. Even taking the reported small anatomical offset of about 1° between these cells into account [4] did not compensate for the small temporal differences but actually resulted in even worse directional tuning. This indicates that sizable differences not only in the temporal but also in the spatial domain are a prerequisite for direction selectivity. Similar considerations are true for all combinations that place neurons with similar response kinetics on neighboring arms. In general, the most effective hybrid detectors result from combinations of cells that are arranged such that the respective HR and BL sub-detectors are as direction selective as possible and aligned in their PD.

For detectors where two low-pass filters flank a central band-pass filter element, both Mi1 and Tm3 seem feasible to fill the role of the latter. However, a previous study blocking the synaptic output of Tm3 found an effect on the response of tangential cells to moving ON edges only at high but not at low to moderate velocities [21]. Hence, although we do not exclude a functional role for Tm3 in ON motion detection, this finding argues against Tm3 as the (sole) central arm of the detector in the ON pathway, as the interference especially with the central arm should fully abolish the detection of motion.

Taken together, an implementation of the ON elementary motion detector as depicted in Figure 6 seems most likely: Mi1 as the fast central input, flanked by the low-pass elements Mi4 and Mi9 constituting the suppressing and enhancing arm in either order. Depending on the location, these neurons need to be either both excitatory or both inhibitory to accommodate their respective response polarity and fulfill the required role of enhancing and suppressing input. Considering their opposite polarity and reciprocal connection, it is more likely that both neurons are inhibitory. This would place Mi9 on the enhancing arm ("A" in Figure 6A), and Mi4 on the suppressing arm ("C" in Figure 6A). Importantly, with the observed range of temporal response characteristics in the input elements, it is not necessary to postulate further delays at the synaptic or dendritic level.

In the OFF pathway, the algorithmic structure of motion detection is less clear. On the one hand, spatiotemporal receptive field measurements of T5 neurons reveal excitatory and inhibitory sub-fields that are offset along the PD axis and appropriately tilted in space and time to support PD enhancement and ND suppression [23]. This would suggest a similar architecture as for T4. On the other hand, other studies have only reported PD enhancement for T5 [22, 39]. Nevertheless, we performed analogous simulations based on the measured T5 input kinetics and receptive fields assuming a similar detector architecture. The two detectors with the highest direction selectivity incorporated the low-pass filter, Tm9, into their suppressing arm (Figure 5E). Lacking a second pure low-pass filter input in the OFF pathway, the central and enhancing arms were occupied with band-pass filters. Because the PDs of PD enhancement and ND suppression need to be aligned, the fastest element of the combination, principally Tm2, must be located in the central position. This is also illustrated by the two worst combinations (Figure 5E, right), where even though the BL module on its own performs quite well, the oppositely oriented HR module destroys the direction selectivity of the hybrid detector. As above, hybrid detectors with the low-pass filter, Tm9, on the central arm perform poorly, as the constituting BL half contributes little to direction selectivity in those combinations (Figure 5E, right; e.g., Tm2xTm9/Tm4).

According to our simulations, and if the structure for T5 resembles the hybrid detector proposed for T4, the arrangement of a central Tm2, flanked on the null side by an enhancing Tm1 and on the preferred side by a suppressing Tm9 input, achieves by far the best direction selectivity. This implementation would predict inhibitory/suppressing input from Tm9 onto T5, which could be experimentally tested. Consistent with this arrangement, out of all four T5 columnar inputs, blocking the synaptic output from Tm4 cells results in the lowest reduction in OFF-edge responses in tangential cells [20]. Nevertheless, those blocking experiments indicate that Tm4 plays a role in the detector that awaits resolving.

In this study, we have shown that the activation of the octopamine system by CDM shifts the temporal-frequency and velocity tuning of T4 and T5 neurons to higher temporal frequencies/velocities, mirroring the temporal tuning shift in tangential cells of the lobula plate observed under active locomotion or octopaminergic activation [24–27]. At the level of T4 and T5 neurons, we find a temporal-frequency optimum of about 1–1.5 Hz for moving gratings under control conditions, corroborating previous studies [14, 17, 22]. Application of CDM shifts the temporal-frequency optimum to 2.5 Hz. T4 and T5 cells show a velocity optimum for moving edges at 3°/s or lower under control conditions, which shifts to about 12°/s under CDM. The much higher velocity optimum observed in tangential cells [40] results from the summation of synaptic inputs from the larger number of T4 and T5 neurons swept by the edge during the same time interval at higher velocities.

In parallel to the temporal-frequency tuning shift in T4 and T5 neurons, the temporal response properties of the input elements, in particular of the band-pass filter elements, accelerate. Indeed, the shift in the tuning of T4 and T5 neurons (Figure 3) can be fully accounted for by the speeding of their input elements (Figures 4 and 5). This further supports the hypothesis that the temporal kinetics of the input elements alone, without any further filtering at the synaptic or T4/T5 dendritic levels, represent the delay stage of the elementary motion detectors.

Interestingly, we observe that whereas the order of input elements with respect to their filter characteristics generally remained the same under CDM, Tm1 became faster than Tm4 (Figure S6). As a consequence, simulated motion detectors using combinations that relied on temporal differences between these two cell types suffered a strong reduction or complete breakdown of direction selectivity under CDM (Figure 5). Considering cell-to-cell variability and such changes under different physiological conditions, detectors relying on small differences in the dynamics of their input elements [18, 39] will not be robust.

Octopaminergic neurons broadly innervate the optic lobes, specifically the medulla, lobula, and lobula plate [26, 28, 29]. They activate during flight and are necessary and sufficient for the observed change in the temporal tuning profile of tangential cells [26]. Although the molecular and cellular mechanisms of action on the medulla neurons and T4/T5 cells, as well as the precise physiological activation of the octopamine system, are beyond the scope of this study, a few points are worth noting. Four different types of octopamine receptors exist in *Drosophila* that are all G protein-coupled receptors but act via different pathways and thus will have different effects [41, 42]. Of those four types, only the octopamine receptors Oamb and to a lesser degree Oct1βR appear to be expressed in the optic lobes [42]. The expression pattern of these octopamine receptors is not known at the cellular level. Considering that all input elements in both the ON and the OFF pathway are accelerated in their responses, albeit to different degrees, it is entirely possible that those changes are indirect and inherited from neurons in the lamina or even the retina. For example, an accelerated response in L1 and L2, and to a smaller degree in L3, could explain the observed response changes in the medulla neurons described here. So far, octopaminergic neurons have not been shown to innervate the retina and lamina directly [26, 28, 29], yet octopamine might nevertheless directly or indirectly affect photoreceptors or lamina neurons. For example, lamina wide-field neurons, projecting from the medulla back into the lamina and forming synaptic inputs to lamina neurons [43], are modulated by the behavioral state and octopamine signaling [44]. Although it cannot be excluded that octopamine acts at multiple levels, including on T4/T5 neurons directly, we have shown that the observed tuning shift in T4/T5 neurons can be fully accounted for by the changes in the temporal dynamics of their input elements.

Pharmacological activation, like any optogenetic or other exogenous activation of the octopamine system, is unlikely to capture all subtleties of the physiological changes during active locomotion, yet it can serve as a tool to manipulate the tuning of the visual motion detection system. At the same time, considering the match between pharmacological manipulation and physiological state changes observed at the level of lobula plate tangential cells [24–26], it is highly likely that the speeding of the filter characteristics in the medulla neurons described here is relevant under physiological conditions.

We have shown that it is possible to construct a hybrid HR/BL detector (as proposed in [17]) with the measured filters for the cellular elements for both the ON and the OFF pathway across different network states. From these, we can predict anatomical arrangements that would give rise to the observed response characteristics of the elementary motion detectors. Although we cannot rule out additional synaptic or dendritic filter mechanisms, we show that the temporal dynamics of the input elements alone are sufficient to explain the response properties of the elementary motion detectors across different tuning regimes. Future studies using the genetic toolbox of *Drosophila* to activate or block individual input neurons and studying the effects on visual responses in the T4 and T5 cells, as well as neurotransmitter and receptor expression pattern analyses and electron-microscopic reconstructions of the wiring, will be required to verify and further confine the proposed circuitry.

EXPERIMENTAL PROCEDURES

Experimental procedures are described in detail in the Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2017.01.051.

AUTHOR CONTRIBUTIONS

A.A., M.S.D., and A.B. conceived the study and designed the experiments. A.A. conducted and analyzed the measurements of T4/T5 cell responses. M.S.D. designed the projector-based stimulation arena and performed and analyzed the measurements of the OFF-pathway elements. F.G.R. performed and analyzed the experiments describing the ON-pathway neurons. G.A. performed and analyzed the patch-clamp recordings from lobula plate tangential cells. M.S.D. performed the computer simulations. A.A. wrote the manuscript with the help of all authors.

ACKNOWLEDGMENTS

We thank Stefan Prech for help in designing the projector-based stimulus arena, Jürgen Haag for sharing his expertise in two-photon microscopy, and

Wolfgang Essbauer and Michael Sauter for fly husbandry. We are very grateful to Michael Reiser, Gerry Rubin, and Aljoscha Nern (Janelia) for sharing unpublished fly lines. We would also like to thank Aljoscha Leonhardt for many helpful discussions, and Alex Mauss for critically reading the manuscript. This work was supported by an EMBO Long-Term Fellowship (A.A.), the Deutsche Forschungsgemeinschaft (SFB 870), and the Max Planck Society.

Received: October 28, 2016
Revised: December 19, 2016
Accepted: January 25, 2017
Published: March 23, 2017

REFERENCES

1. Borst, A. (2014). Fly visual course control: behaviour, algorithms and circuits. Nat. Rev. Neurosci. *15*, 590–599.
2. Fischbach, K.F., and Dittrich, A.P.M. (1989). The optic lobe of *Drosophila melanogaster*. I. A Golgi analysis of wild-type structure. Cell Tissue Res. *258*, 441–475.
3. Bausenwein, B., Dittrich, A.P., and Fischbach, K.F. (1992). The optic lobe of *Drosophila melanogaster*. II. Sorting of retinotopic pathways in the medulla. Cell Tissue Res. *267*, 17–28.
4. Takemura, S.-Y., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P.K., Katz, W.T., Olbris, D.J., Plaza, S.M., Winston, P., et al. (2013). A visual motion detection circuit suggested by *Drosophila* connectomics. Nature *500*, 175–181.
5. Shinomiya, K., Karuppudurai, T., Lin, T.-Y., Lu, Z., Lee, C.-H., and Meinertzhagen, I.A. (2014). Candidate neural substrates for off-edge motion detection in *Drosophila*. Curr. Biol. *24*, 1062–1070.
6. Barlow, H.B., and Levick, W.R. (1965). The mechanism of directionally selective units in rabbit's retina. J. Physiol. *178*, 477–504.
7. 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.
8. Borst, A., and Helmstaedter, M. (2015). Common circuit design in fly and mammalian motion vision. Nat. Neurosci. *18*, 1067–1076.
9. 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.
10. 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.
11. 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.
12. 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.
13. Joesch, M., Weber, F., Eichner, H., and Borst, A. (2013). Functional specialization of parallel motion detection circuits in the fly. J. Neurosci. *33*, 902–905.
14. 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.
15. 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.
16. 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.
17. Haag, J., Arenz, A., Serbe, E., Gabbiani, F., and Borst, A. (2016). Complementary mechanisms create direction selectivity in the fly. eLife *5*, e17421.
18. 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.
19. Behnia, R., and Desplan, C. (2015). Visual circuits in flies: beginning to see the whole picture. Curr. Opin. Neurobiol. *34*, 125–132.
20. 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.
21. 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.
22. Fisher, Y.E., Silies, M., and Clandinin, T.R. (2015). Orientation selectivity sharpens motion detection in *Drosophila*. Neuron *88*, 390–402.
23. 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.
24. 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.
25. Jung, S.N., Borst, A., and Haag, J. (2011). Flight activity alters velocity tuning of fly motion-sensitive neurons. J. Neurosci. *31*, 9231–9237.
26. 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.
27. Longden, K.D., and Krapp, H.G. (2010). Octopaminergic modulation of temporal frequency coding in an identified optic flow-processing interneuron. Front. Syst. Neurosci. *4*, 153.
28. Sinakevitch, I., and Strausfeld, N.J. (2006). Comparison of octopamine-like immunoreactivity in the brains of the fruit fly and blow fly. J. Comp. Neurol. *494*, 460–475.
29. Busch, S., Selcho, M., Ito, K., and Tanimoto, H. (2009). A map of octopaminergic neurons in the *Drosophila* brain. J. Comp. Neurol. *513*, 643–667.
30. 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.
31. Roeder, T. (1995). Pharmacology of the octopamine receptor from locust central nervous tissue (OAR3). Br. J. Pharmacol. *114*, 210–216.
32. Schnell, B., Weir, P.T., Roth, E., Fairhall, A.L., and Dickinson, M.H. (2014). Cellular mechanisms for integral feedback in visually guided behavior. Proc. Natl. Acad. Sci. USA *111*, 5700–5705.
33. Mazurek, M., Kager, M., and Van Hooser, S.D. (2014). Robust quantification of orientation selectivity and direction selectivity. Front. Neural Circuits *8*, 92.
34. 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.
35. Fisher, Y.E., Leong, J.C.S., Sporar, K., Ketkar, M.D., Gohl, D.M., Clandinin, T.R., and Silies, M. (2015). A class of visual neurons with wide-field properties is required for local motion detection. Curr. Biol. *25*, 3178–3189.
36. Yang, H.H., St-Pierre, F., Sun, X., Ding, X., Lin, M.Z., and Clandinin, T.R. (2016). Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. Cell *166*, 245–257.
37. Takemura, S.-Y., Xu, C.S., Lu, Z., Rivlin, P.K., Parag, T., Olbris, D.J., Plaza, S., Zhao, T., Katz, W.T., Umayam, L., et al. (2015). Synaptic circuits and their variations within different columns in the visual system of *Drosophila*. Proc. Natl. Acad. Sci. USA *112*, 13711–13716.
38. 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.

39. 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.

40. 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.

41. Farooqui, T. (2007). Octopamine-mediated neuromodulation of insect senses. Neurochem. Res. *32*, 1511–1529.

42. El-Kholy, S., Stephano, F., Li, Y., Bhandari, A., Fink, C., and Roeder, T. (2015). Expression analysis of octopamine and tyramine receptors in *Drosophila*. Cell Tissue Res. *361*, 669–684.

43. Rivera-Alba, M., Vitaladevuni, S.N., Mishchenko, Y., Lu, Z., Takemura, S.Y., Scheffer, L., Meinertzhagen, I.A., Chklovskii, D.B., and de Polavieja, G.G. (2011). Wiring economy and volume exclusion determine neuronal placement in the *Drosophila* brain. Curr. Biol. *21*, 2000–2005.

44. 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.

Current Biology, Volume *27*

Supplemental Information

The Temporal Tuning of the *Drosophila* Motion Detectors Is Determined by the Dynamics of Their Input Elements

Alexander Arenz, Michael S. Drews, Florian G. Richter, Georg Ammer, and Alexander Borst

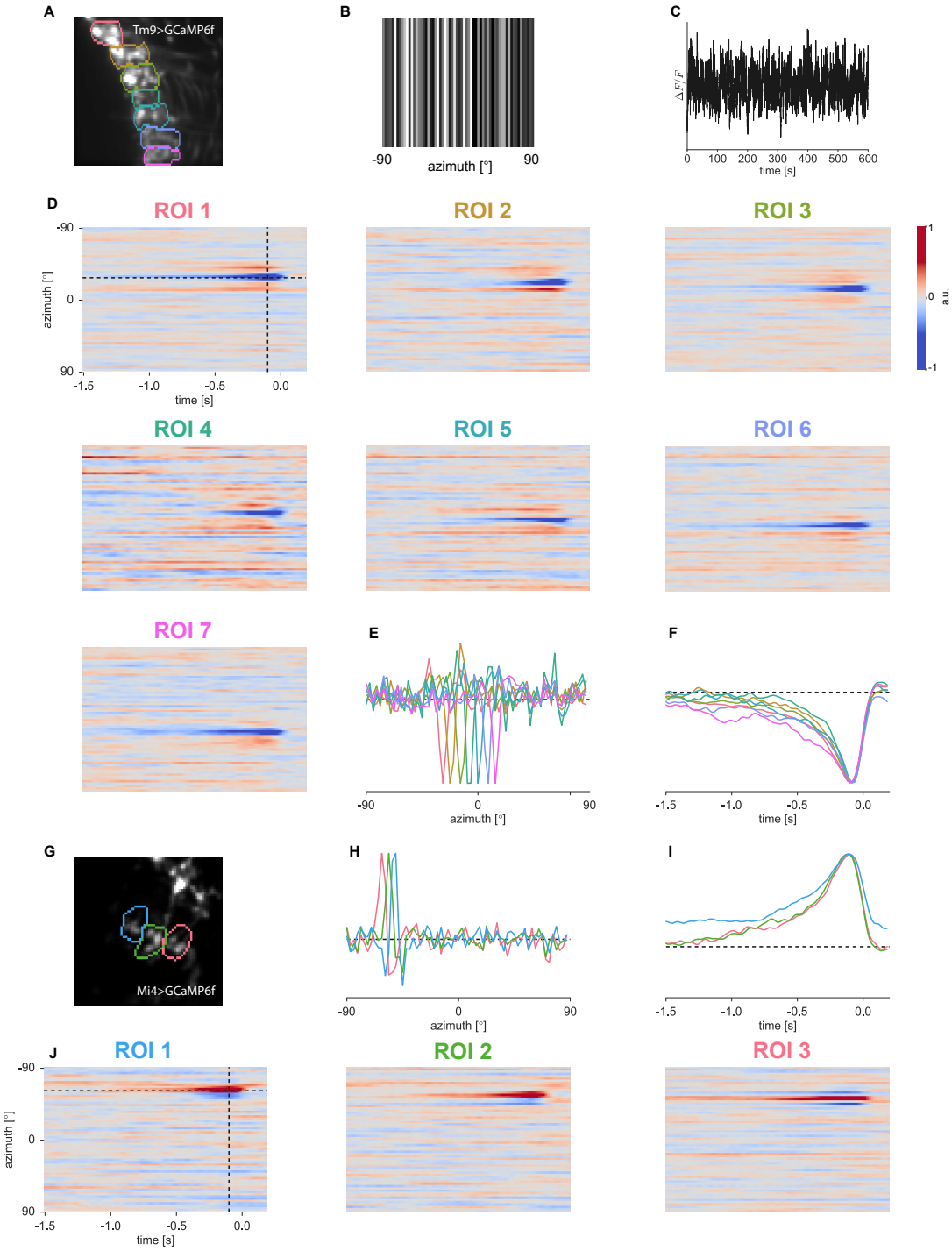

A
Tm9>GCaMP6f
B
-90
azimuth [°]
90
C
ΔF/F
0 100 200 300 400 500 600
time [s]
D
ROI 1
ROI 2
ROI 3
ROI 4
ROI 5
ROI 6
ROI 7
azimuth [°]
-90
0
90
-1.5 -1.0 -0.5 0.0
time [s]
1
0
-1
a.u.
E
-90 0 90
azimuth [°]
F
-1.5 -1.0 -0.5 0.0
time [s]
G
Mi4>GCaMP6f
H
-90 0 90
azimuth [°]
I
-1.5 -1.0 -0.5 0.0
time [s]
J
ROI 1
ROI 2
ROI 3
azimuth [°]
-90
0
90
-1.5 -1.0 -0.5 0.0
time [s]

Figure S1. Related to Figure 2. Example acquisition of spatiotemporal receptive fields via stochastic stimulation and reverse correlation of calcium signals for the neurons Tm9 and Mi4.

(A) 2-photon image from a fly expressing GCaMP6f in Tm9 axon terminals in the lobula. Highlighted in color are seven manually drawn regions of interest (ROIs) around individual terminals from neighboring columns.

(B) Snapshot of one frame of the one-dimensional horizontal noise stimulus.

(C) Calcium trace from a single ROI in response to 10 minutes of white noise stimulation.

(D) Spatiotemporal receptive fields obtained by reverse correlation of the calcium signals in each ROI with the stimulus.

(E) Cross-sections through the receptive fields along the space axis reflecting the retinotopic organisation of the lobula.

(F) Cross-sections through the receptive fields along the temporal axis revealing the low-pass characteristics of Tm9.

(G-J) Same for Mi4.

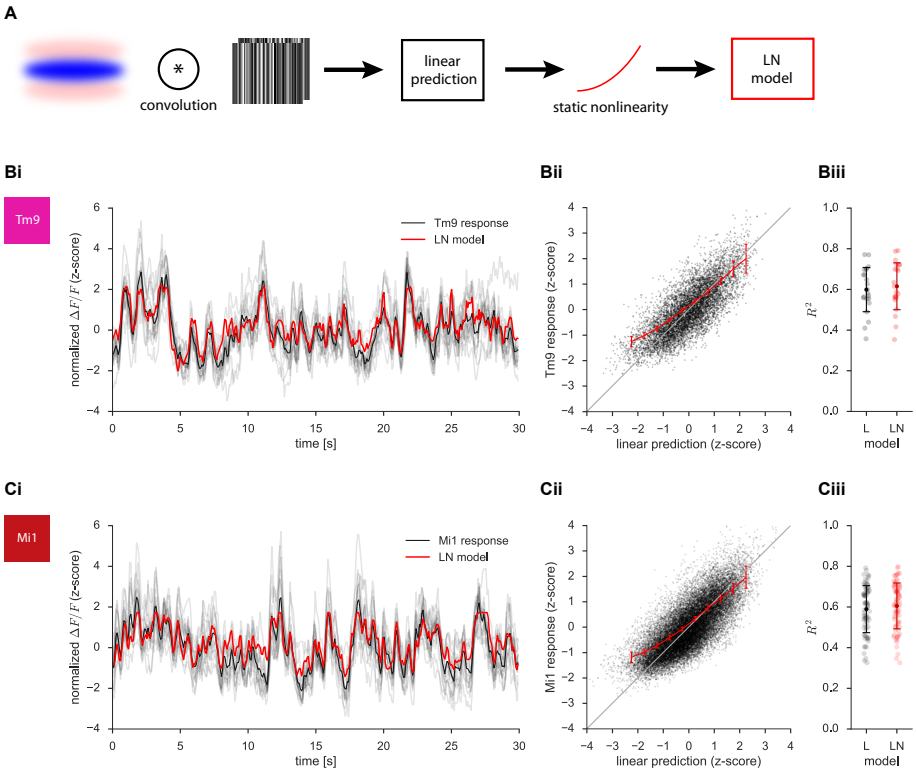


Figure S2. Related to Figure 2. Prediction of calcium responses in Tm9 and Mi1 terminals from the linear spatiotemporal receptive fields.
(A) Schematic of the model. The linear prediction of individual axon terminal responses (of Tm9 or Mi1) to a white noise stimulus is given by the convolution of the stimulus with the respective spatio-temporal receptive field of the cell. A linear-nonlinear model (LN) is built by remapping the output of the linear prediction with a static nonlinearity.
(Bi) Actual response of an exemplary Tm9 axon terminal (black) and the prediction of the LN model (red). (Bii) Scatter plot of the linear prediction against the actual response for all cells recorded. The static nonlinearity (red) is obtained by averaging the point cloud within discrete bins along the x-axis for each axon terminal. (Biii) Coefficient of determination for the linear model (L, black) and the linear-nonlinear model (LN, red). The linear model prediction alone accounted for 60% and the LN model for 62% of the response variance. Circles represent measurements of individual terminals, the bar shows the standard deviation and the mean among all cells measured (N = 4, n = 22).
(C) Same as in (B), but for Mi1 (N = 4, n = 78). The L model alone accounted for 59% and the LN model for 61% of the response variance.

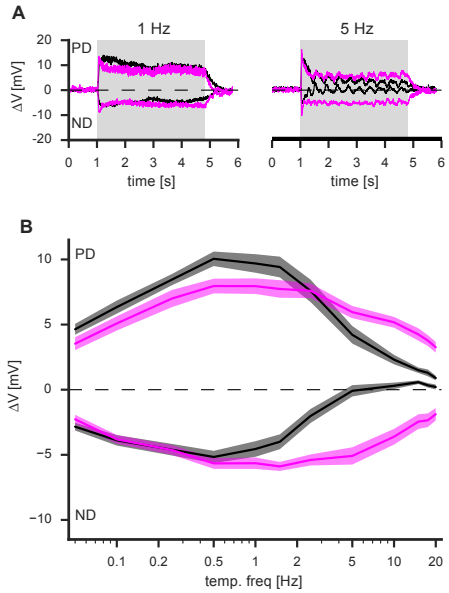
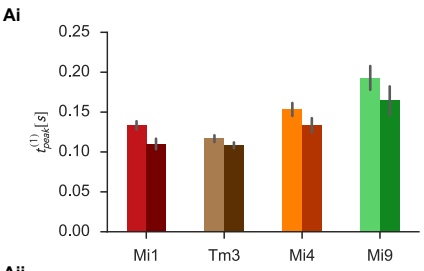
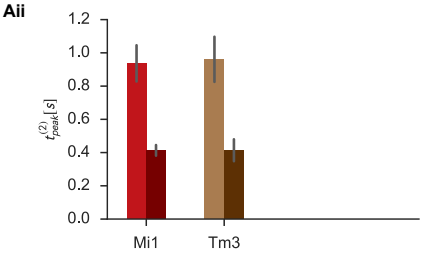


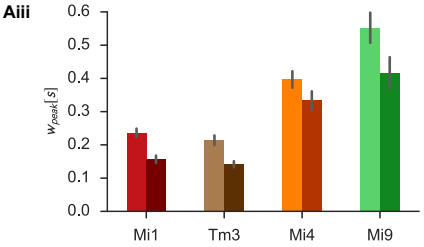
Figure S3. Related to Figure 3. Temporal frequency tunings of lobula plate tangential cells change with the application of the octopamine agonist CDM.

(A) Voltage responses of HS and VS tangential cells in the lobula plate (population average, N = 15 flies, n=15 cells) to square-wave gratings moving in the preferred or null direction in control (black) and after application of CDM (magenta) for gratings moving at a temporal frequency of 1 Hz (left) or 5 Hz (right). The period of motion of the grating is indicated by the grey-shaded region.

(B) Average voltage responses over the stimulation period for square-wave gratings at different temporal frequencies. Responses, measured as average voltage deflections over the whole stimulus period, peaked at 0.5 Hz in both the preferred (as maximum average depolarization) and null (as maximum average hyperpolarization) direction. Application of the octopamine agonist chlordimeform (CDM; magenta) at a final concentration of 20 µM resulted in increased responses to higher temporal frequencies from 2-20 Hz.



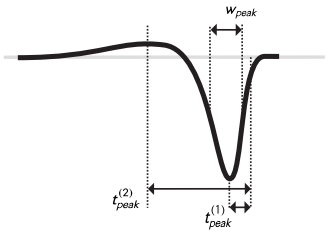


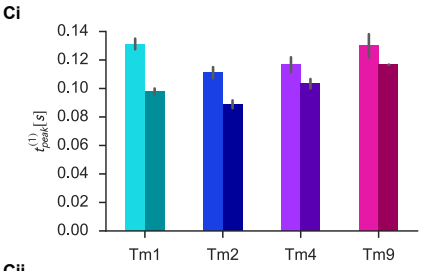


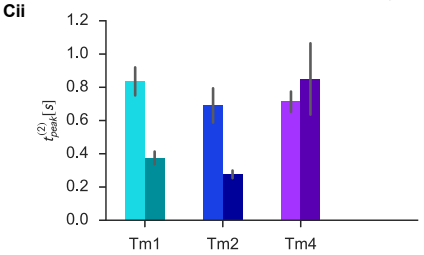
Aiv

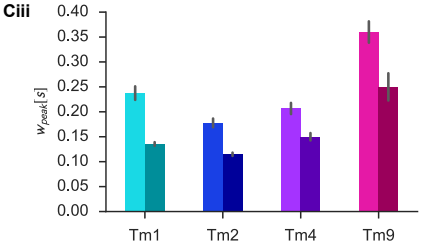
	$t^{(1)}_{peak}$	$t^{(2)}_{peak}$	w_{peak}
Mi1	0.108	0.013	0.014
Tm3	0.076	0.004	< 0.001
Mi4	0.033	n.a	0.041
Mi9	0.259	n.a	0.037

control vs CDM









Civ

	$t^{(1)}_{peak}$	$t^{(2)}_{peak}$	$w_{peak}[s]$
Tm1	< 0.001	0.004	< 0.001
Tm2	< 0.001	0.005	< 0.001
Tm4	0.016	0.564	0.003
Tm9	0.178	n.a	0.015

control vs CDM

Figure S4. Related to Figure 2 and 4. Analysis of the temporal filters of the inputs to T4 and T5.

(A) T4 inputs. We quantified the shape of the temporal filters, as well as their change after application of CDM, by three measures (B): the time-to-peak for the first peak ($t^{(1)}_{peak}$, (Ai)) and, for the biphasic filter kernels of band-pass filters, to the second peak ($t^{(2)}_{peak}$, (Aii)), as well as the full-width at half-maximum for the first peak (w_{peak}, (Aiii)) of the temporal kernel. Measurements after application of CDM are presented in a darker color shade (right bars) than for the control condition (left bars) of the respective cell. Statistical comparisons between control and CDM condition (based on a paired t-test) are shown in (Aiv), highlighted in red color when statistical significance is observed (with a gradient in the red nuance from light to dark red indicating p values of $p<0.05$, $p<0.01$ and $p<0.001$).

(C) Same, for T5 inputs.

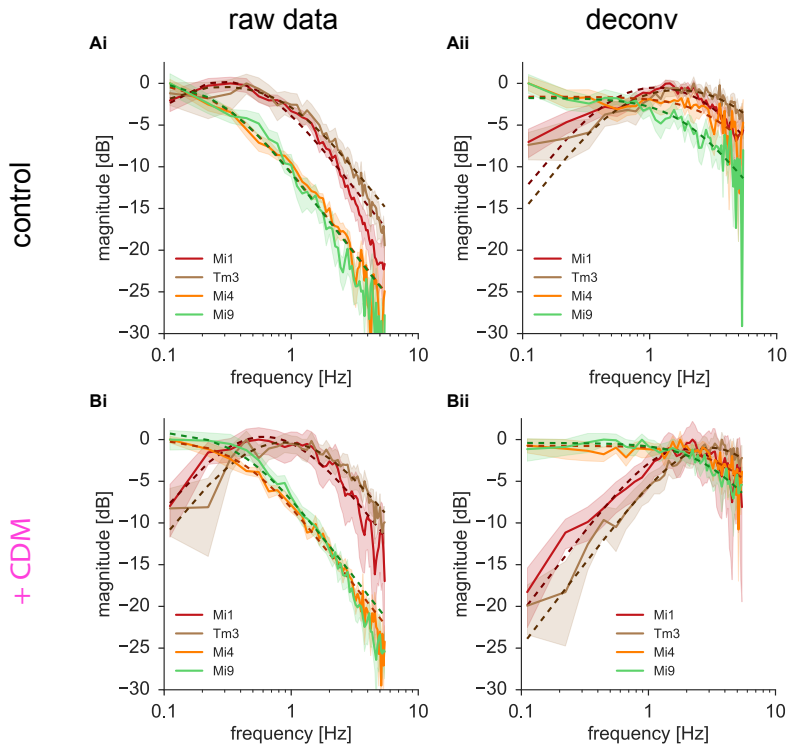


Figure S5. Related to Figure 5. Model fitting on the frequency spectra of the ON pathway elements.
(Ai) Frequency spectrum derived from calcium imaging experiments for the ON pathway columnar neurons Mi1, Tm3, Mi4, Mi9. (Aii) Frequency spectrum after deconvolution with a low-pass filter representing the dynamics of the calcium indicator GCaMP6f. Dashed lines represent the fitted frequency responses of 1st order band-pass or low-pass filters.
(Bi, Bii) Like (Ai, Aii), for the spectra determined from the recordings after application of CDM.

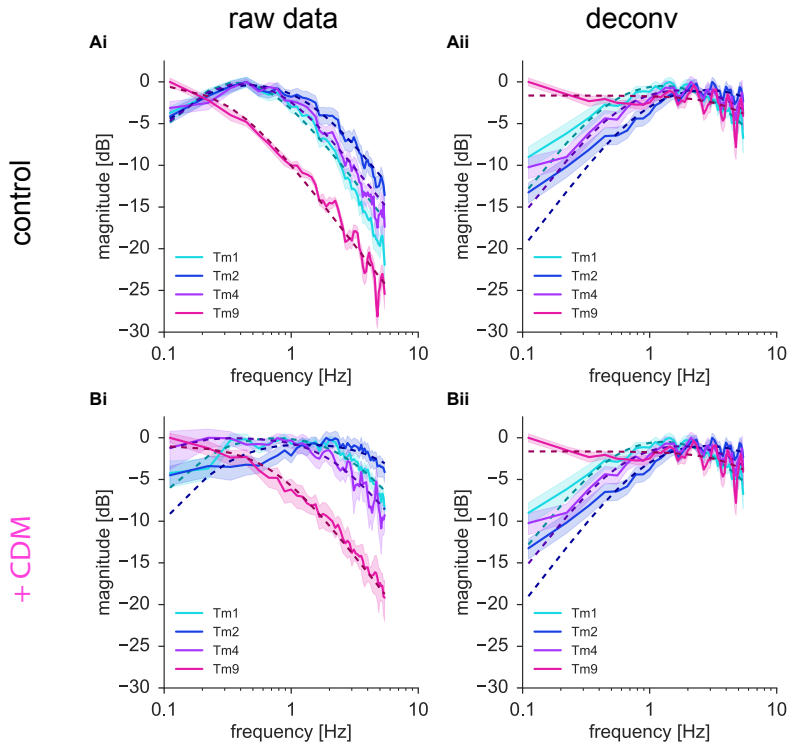


Figure S6. Related to Figure 5. Model fitting on the temporal filter frequency spectra of the OFF pathway elements.

(Ai) Measured frequency spectra based on calcium imaging experiments for the OFF pathway elements Tm1, Tm2, Tm4 and Tm9. (Aii) Frequency spectra after deconvolution with a filter describing the dynamics of the calcium indicator. Dashed lines represent the fitted frequency responses of 1st order band-pass or low-pass filters.

(Bi, Bii) Same as (Ai, Aii), after application of CDM.

	Mi1		Tm3		Mi4		Mi9	
	control	+ CDM	control	+ CDM	control	+ CDM	control	+ CDM
A_{raw}	1.271	1.314	0.996	1.204	1.007	1.000	0.982	1.060
$t_{HP, raw}$	1.078 s	0.445 s	1.769 s	0.340 s				
$t_{LP, raw}$	0.266 s	0.143 s	0.158 s	0.086 s	0.519 s	0.370 s	0.546 s	0.373 s
$R^2_{t, raw}$	0.975	0.942	0.971	0.925	0.984	0.990	0.982	0.987
A_{deconv}	1.146	1.643	1.035	2.034	0.831	0.906	0.789	0.940
$t_{HP, deconv}$	0.318 s	0.075 s	0.260 s	0.044 s				
$t_{LP, deconv}$	0.054 s	0.075 s	0.027 s	0.044 s	0.038 s	0.028 s	0.077 s	0.043 s
$R^2_{t, deconv}$	0.872	0.851	0.768	0.955	0.666	0.643	0.832	0.906
A_{sur} / A_{cen}	0.022	0.028	0.000	0.000	0.132	0.093	0.063	0.043
$FWHM_{cen}$	6.81 °	7.09 °	11.91 °	12.40 °	6.47 °	6.49 °	6.37 °	6.37 °
$FWHM_{sur}$	28.81 °	19.86 °	-	-	16.14 °	20.02 °	23.98 °	28.08 °
R^2_{s}	0.995	0.990	0.985	0.984	0.987	0.989	0.980	0.992

Table S1. Related to Figure 2 and 4. Spatio-temporal response properties of T4 input cells.
Numerical parameters derived from the model fits to the temporal and the spatial components of the spatio-temporal receptive fields obtained from reverse correlation for the ON-pathway neurons.

	Tm1		Tm2		Tm4		Tm9	
	control	+ CDM	control	+ CDM	control	+ CDM	control	+ CDM
A_{raw}	1.419	1.072	1.068	0.933	1.218	0.990	0.961	0.890
$t_{HP, raw}$	0.632 s	0.754 s	0.962 s	0.558 s	0.788 s	2.210 s		
$t_{LP, raw}$	0.271 s	0.051 s	0.113 s	0.020 s	0.186 s	0.070 s	0.462 s	0.220 s
$R^2_{s, raw}$	0.978	0.901	0.979	0.564	0.972	0.930	0.985	0.981
A_{deconv}	1.117	2.217	1.038	2.657	1.018	1.002	0.827	0.836
$t_{HP, deconv}$	0.296 s	0.032 s	0.153 s	0.018 s	0.249 s	0.086 s		
$t_{LP, deconv}$	0.044 s	0.032 s	0.014 s	0.018 s	0.024 s	0.015 s	0.017 s	0.008 s
$R^2_{s, deconv}$	0.895	0.974	0.893	0.990	0.807	0.880	0.273	0.226
A_{sur} / A_{cen}	0.040	0.029	0.035	0.026	0.054	0.005	0.046	0.071
$FWHM_{cen}$	8.12 °	7.34 °	7.93 °	7.23 °	11.45 °	10.10 °	6.92 °	6.53 °
$FWHM_{sur}$	27.14 °	20.46 °	30.52 °	23.01 °	34.62 °	72.03 °	23.78 °	21.74 °
R^2_s	0.997	0.995	0.994	0.991	0.992	0.988	0.995	0.995

Table S2. Related to Figure 2 and 4. Spatio-temporal response properties of T5 input cells.
Numerical parameters derived from the model fits to the temporal and the spatial components of the spatio-temporal receptive fields obtained from reverse correlation for the OFF-pathway neurons.

Supplemental Experimental Procedures

Flies/preparation

Flies were raised and kept on standard cornmeal-agar medium on a 12 hour light/12 hour dark cycle at 25°C and 60% humidity. For patch-clamp recordings from tangential cells, Canton S flies were used. For calcium imaging experiments, the genetically-encoded calcium indicators GCaMP6f or GCaMP6m [S1] were expressed using the Gal4/UAS- or LexA/lexAop-system in cell-type specific driver lines, resulting in the following genotypes:

Short name	**Genotype**
Mi1>GC6f	*w-; R19F01-AD/UAS-GCaMP6f; R71D01-DBD/UAS-GCaMP6f*
Tm3>GC6f	*w-; UAS-GCaMP6f; R13E12-Gal4*
Mi4>GC6f	*w-; R48A07-AD/UAS-GCaMP6f; R13F11-DBD/UAS-GCaMP6f*
Mi9>GC6f	*w-; R48A07-AD/UAS-GCaMP6f; VT046779-DBD/UAS-GCaMP6f*
Tm1>GC6f	*w-; UAS-GCaMP6f; VT12717-Gal4*
Tm2>GC6f	*w-; UAS-GCaMP6f; VT12282-Gal4*
Tm4>GC6f	*w-; UAS-GCaMP6f; R35H01-Gal4*
Tm9>GC6f	*w-; UAS-GCaMP6f; VT65303-Gal4*
T4/T5>GC6m	*w-; Sp/CyO ; VT50384-lexA, lexAop-GCaMP6m/TM6b*

The transgenic fly lines driving split-Gal4 expression in the medulla neurons Mi1, Mi4 and Mi9, respectively, were generated and will be described in [S2] (with the Mi1 driver line corresponding to their transgenic fly line SS00809, Mi4 to SS01019, and Mi9 to SS02432).

For electrophysiological and calcium imaging experiments, flies were prepared as previously described [S3, S4]. Briefly, flies were anaesthetized on ice or with CO_2, fixed with their backs, legs and wings to a Plexiglas holder with the back of the head exposed to a recording chamber filled with fly external solution. The cuticula at the back of the head on one side was cut away with a fine hypodermic needle and removed together with muscles and air sacks covering the underlying optic lobe. To gain access to tangential cells for electrophysiological recordings, the neurolemma covering the brain was partially digested by applying 0.5mg/ml collagenase IV (Gibco) with a glass electrode to the brain until the tangential cell somata were exposed. Where indicated, the octopamine agonist chlordimeform (CDM, Sigma Aldrich) was added as a 2mM stock solution (in external solution) directly to the bath to yield a final concentration of 20 μM. Diffusion was allowed for 15 min before recordings recommenced.

Patch-clamp recordings from vertical and horizontal system tangential cells were performed as previously described [S4].

2-Photon calcium imaging

Calcium imaging was performed on custom-built 2-photon microscopes as previously described [S3] controlled with the ScanImage software in Matlab [S5]. Acquisition rates were between 3.8 and 15 Hz, image resolution between 64x64 and 128x128 pixels. Before starting the acquisition, we verified that the receptive fields of the cells were located on the stimulus arena by showing a search stimulus consisting of moving gratings.

Visual stimulation

For the study of visual responses of lobula plate tangential cells and T4/T5 neurons, visual stimuli were presented on an **LED arena**, based on a design by [S6], covering approximately 180° in azimuth and 90° in elevation. Stimuli covered the whole extent of the arena and were presented at full contrast. Square-wave gratings had a spatial wavelength of 24°, and moved with velocities of 1.2-480°/s in the preferred and null direction, corresponding to temporal frequencies ranging from 0.05 to 20°/s. Single stimulation periods of moving gratings lasted for 3.8 s, separated by periods of 5 s where the grating remained stationary. For the edge velocity tuning, bright or dark edges of full contrast were presented, moving at velocities of 3 to 300 °/s in the preferred (up) and null direction (down) of T4c/T5c neurons, separated by 6 s. All stimuli were presented in a pseudo-random order with 3-5 repetitions per stimulus.

The spatio-temporal response properties of the Mi and Tm columnar input elements were determined on a custom-built **projector-based arena** that allowed for greater stimulus flexibility. Stimuli were projected with 2 commercial micro-projectors (TI DLP Lightcrafter 3000) onto the back of an opaque cylindrical screen covering 180° in azimuth and 105° in elevation of the fly's visual field. The projectors were programmed to use only the green LED (OSRAM L CG H9RN) which emits light between 500nm to 600nm wavelength. This increased the refresh rate from 60 to 180 Hz (at 8 bit color depth). To prevent overlap between the spectra of the GCaMP signal and the arena light, we placed two long-pass filters (Thorlabs FEL0550 and FGL550) in front of each projector restricting the stimulus light to wavelengths above 550nm. A band-pass filter in front of the photomultiplier (Brightline 520/35) allowed only the portion of the light within the GCaMP emission spectrum to be detected. Additional shielding of stray light from the arena with black foil effectively suppressed any leak of the arena light into the photomultiplier signal. The maximum luminance achieved by our stimulation system is $276 \pm 48\ \mathrm{cd/m^2}$. For all stimuli used here, we set the medium brightness to a 8-bit grayscale value of 50, which corresponds to a medium luminance of $55 \pm 11\ \mathrm{cd/m^2}$.

Stimuli were rendered using a custom written software in Python 2.7. To account for the curvature of the arena screen, our software pre-distorts the generated images such that the projected image appears as a regular grating on the screen. For that, the software takes advantage of functions from Panda3D, a framework for 3D rendering for Python.

Gaussian noise stimulus

To generate the horizontal white noise stimulus, we partitioned the cylindrical screen into 64 bars, so that each bar covered an angle of approximately 2.8° in azimuth. For each bar, samples were drawn at a frame rate of 60 Hz from a Gaussian distribution, so that the standard deviation was at 25% contrast around a mean intensity value of 50 on the 8-bit grayscale of the display devices. We then filtered the random samples for each bar with a Gaussian filter with a standard deviation of 5 Hz in the frequency domain which leads to a stimulus auto-correlation function that is a Gaussian with approximately 45ms standard deviation. Since the calcium indicator dynamics of GCaMP and the data acquisition frame rate (12 Hz in this case) place a lower bound on the temporal precision of the signal we can extract from calcium imaging experiments, we restricted the frequency content of the stimulus in this way to the relevant domain. The whole stimulus sequence was 10 minutes long and was exported as a video file in H.264 format with lossless compression. For the vertical noise the same stimulus was rotated by 90° and scaled such that 54 bars covered the height of the screen, accounting for the aspect ratio of the screen being approximately 1.2.

Data acquisition and analysis

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

For the **electrophysiological experiments**, baseline-subtracted voltage responses of tangential cells were averaged across trials, and the response to gratings was quantified as the average voltage over the whole period of the respective stimulus presentation. Preferred direction was front-to-back for HS and down for VS cells, null direction the corresponding opposite direction. For Suppl. Figure S3 voltage responses over the individual stimuli were averaged across all cells.

Calcium imaging: Images were automatically registered using vertical and horizontal translations to correct for the movement of the brain. Fluorescence changes (ΔF/F values) were then calculated by dividing every registered frame by the average of the registered first 5 images of the recording. Regions of interest (ROIs) were selected on the average raw image by hand: in layer 10 of the medulla for the ON, in the lobula for the OFF pathway elements, outlining single terminals. For T4 and T5 neurons, ROIs were routinely chosen in the lobula plate, encompassing small regions with single to few axon terminals, or selected to cover single neurites between medulla or lobula and lobula plate. For Figure 3D&E, ROIs were drawn in the medulla for T4 and in the lobula for T5 neurons to separate those 2 cell types. Averaging the fluorescence change over this ROI in space resulted in a ΔF/F time course. Neuronal responses were quantified as the maximum ΔF/F value over the stimulation period plus the subsequent 0.5 s, subtracted by the average of the baseline period covering the 2 frames before the respective stimulus onset. To average across cells/ROIs, responses were first normalized to the maximum response of each ROI to the corresponding stimulus set. For edges, normalization was performed separately to ON and OFF stimuli to take any selection bias for T4 or T5 cells within the ROI into account.

White noise reverse-correlation

For the input elements, spatio-temporal receptive fields were calculated following standard reverse-correlation methods (Figure S1) [S7, S8]. First, the mean value was subtracted from the raw signals of single ROIs by using a low-pass filtered version of the signal (Gaussian filter with 120 seconds standard deviation) as a baseline for a ΔF/F-like representation of the signal. This effectively removed slow baseline fluctuations caused by bleaching and very slow changes in the average calcium level from the signals.
We then calculated the stimulus-response reverse correlation function

$$K(x,\tau) = \int_0^T dt\, S(x, t-\tau) \cdot R(t)$$

where S denotes the stimulus and R the response of the neuron.

The resulting spatiotemporal fields were normalized in z-score and as a quality control only receptive fields with peak amplitudes above 10 standard deviations from the mean were taken for further analysis (for Mi9 the threshold was lowered to 7). Cross-sections through the receptive fields along the space axis were fit with a Gaussian function to determine the position of the peak.

Since one imaging frame is built up continuously over one sample time, ROIs lying at different y-coordinates in the image will in fact be imaged at slightly different times. Since the stimulus is presented at a higher frame rate of 60 Hz, this leads to a notable peak shift between the impulse responses of different ROIs. We corrected for this by translating the spatiotemporal receptive field of each ROI by a) the time difference between the start of a frame and the effective sampling point estimated by the y-coordinate of the center of mass of the respective ROI and b) the start time of the white noise stimulus within the very first frame acquired during stimulation.

Spatio-temporal receptive fields resulting from different ROIs (that were retinotopically shifted) were then centered about each other to generate a mean receptive field. To ensure receptive fields of input elements were fully covered, cells with a receptive field center less than 10 pixels (28°) from the edge of the arena were excluded.

Frozen noise

Filter kernels were validated by testing their ability to predict the neuronal responses from the stimulus. For this, neurons were again stimulated with a white-noise stimulus, only this time part of the stimulus consisted of 15 repetitions (each 30 seconds long) of a white-noise sequence ('frozen noise') to eliminate noise in the neuronal responses. As above, spatio-temporal filter kernels were then reconstructed from responses to single repetition stimulus sequences (20 minutes long). Analogously to above, only receptive fields with a peak higher than 20 standard deviations were included for further analysis. Subsequently the averaged response during the held-out test portion of the stimulus was predicted for each recorded cell individually. Linear predictions were obtained by convolution of the spatio-temporal filter kernels with the frozen noise stimulus along the time axis. Filter kernels were thresholded versions of the spatio-temporal receptive fields (all values below 5% of the peak amplitude as well as regions further away than 15° from the receptive field center were set to zero). Both, the predicted response trace and the actual mean response to the frozen noise stimulus, were normalized in z-score in order to make different cells with varying calcium indicator expression levels and therefore different absolute signal values comparable. The static nonlinearity for the LN model was estimated for each cell by averaging all values from the actual mean response corresponding to values of the predicted response within bins of size 0.5 from -2.5 to +2.5 z-score (see scatter plots Bii and Cii in Suppl. Figure S2). Prediction accuracy of the linear filter was assessed through the correlation of the predicted versus actual response of the neuron [S9].

Spatial receptive field model

The one-dimensional spatial receptive fields (Figure 2 E-H and K-N, top and right) are cross-sections through the peak of the spatiotemporal receptive fields along the space axis and are averaged over the 12 samples (200ms) around the peak. For almost all columnar neurons measured we found a small-field, antagonistic center-surround organization of the spatial receptive field using both the horizontal and the vertical white noise stimulus.

Mathematically, receptive fields of this kind can be described as a difference of Gaussians

$$RF_{1D}(\varphi) = e^{-\frac{\varphi^2}{2\sigma_{cen}{}^2}} - A_{rel} \cdot e^{-\frac{\varphi^2}{2\sigma_{sur}{}^2}}$$

without loss of generality for the horizontal one-dimensional receptive field along the azimuth φ. Here, σ_{cen} and σ_{sur} are the standard deviations of center and surround, respectively, and $A_{rel} = A_{sur}/A_{cen}$ denotes the relative strength of the surround in relation to the amplitude of the center Gaussian (which is normalized to 1).

To reconstruct a two-dimensional receptive field from the measured one-dimensional projections, we chose the same mathematical approach as above, only in 2D:

$$RF_{2D}(\varphi,\vartheta) = e^{-\frac{(\varphi^2+\vartheta^2)}{2\sigma_{cen}{}^2}} - A_{rel}\cdot e^{-\frac{(\varphi^2+\vartheta^2)}{2\sigma_{sur}{}^2}}$$

For simplicity, throughout the analysis we used the small-angle approximation $\tan\vartheta \approx \vartheta$ for the vertical axis or the elevation ϑ even if receptive fields span angles larger than 5°. Thus, we neglected perspective distortions induced by the arena screen not being spherical, but cylindrical. Accounting for additional distortions induced by the relative displacement of the fly's body in relation to the elevation of the receptive field on the arena would require even more detailed mathematical description, yet we did not observe any severe irregularities in the spatial receptive fields.

It is important to note that receptive field estimation via a one-dimensional stimulus as performed here yields in fact a projection of the underlying two-dimensional spatial receptive field:

$$RF_{1D}(\varphi) = \int_{-\infty}^{\infty} RF_{2D}(\varphi,\vartheta)d\vartheta$$

Hence, we fitted the above function $RF_{2D}(\varphi,\vartheta)$ such that its projections along the horizontal and vertical axis would agree with the given one-dimensional receptive field projections measured via reverse correlation. The fitting procedure was implemented using standard least-square algorithms (SciPy 0.16.1). The resulting values for A_{rel}, σ_{cen} and σ_{sur} and the corresponding coefficients of the fit are given in Table S1 and S2 for each neuron type.

Temporal filter model

The time-reversed impulse responses shown in Figure 2&4 are cross-sections through the center of the spatiotemporal receptive fields along the time axis and are averaged over the three center pixels. For the frequency domain representations in Figure 2&4, impulse responses were Fourier-transformed, averaged, and the resulting amplitude spectrum (absolute value) was divided by the power spectrum of the stimulus for frequencies below 5.5 Hz (below the Nyquist frequency). This is equivalent to deconvolving the impulse response with the stimulus auto-correlation and thereby correcting for non-white input signals [S7]. All frequency-space-representations are plotted on a double logarithmic scale expressing all signal gains in decibel according to convention in filter theory.

The complicated relationships between calcium, calcium indicator, voltage and neurotransmitter release of a cell render it impossible to precisely characterize each of these aspects having access to only the calcium indicator fluorescence as a read-out. However, we can assume under certain conditions that the calcium indicator itself essentially acts as a simple low-pass filter on the calcium signal [S10], which is a kind of distortion that we are able to correct for by applying deconvolution.

GCaMP6f is designed to have especially fast kinetics. However, we can find decay constants in the order of several hundreds milliseconds that vary depending on the system under observation [S1]. As an approximation we chose a time constant of 350 ms for a plausible low-pass filter that distorted the calcium signals in our system, which lies in the range of reported decay constants for GCaMP6f [S1, S11].

We corrected the frequency domain representations of the temporal filters of all cells by dividing the spectra with the frequency response of a 1st order low-pass filter with this time constant. Since this was restricted to frequencies below the Nyquist frequency, we did not have to apply additional techniques to avoid the impact of poor signal-to-noise ratios at higher frequencies.

For quantitative description and further simulations, we sought to describe the response characteristic of each cell under each condition with a simplified model that catches the main properties. For that, we fitted simple 1st order filters to the corrected frequency responses of all cells. We did this separately for each condition, i.e. for control and CDM condition and for the raw filters (corrected by the stimulus power spectrum only) and the deconvolved filters (corrected by the GCaMP filter) respectively.

In particular, we approximated Mi1, Tm3, Tm1, Tm2 and Tm4 as band-pass filters and fitted a band-pass model consisting of a 1st order high-pass and a 1st order low-pass filter to the frequency responses (Figures S5, S6). The band-pass model was parametrized by a multiplicative amplitude and the two time constants of the filters. Parameters were optimized using a standard implementation of the Levenberg-Marquardt algorithm (SciPy). Similarly, Mi4, Mi9 and Tm9 frequency responses were fit using a 1st order low-pass filter model.

Computational modeling

Neural simulations (Figure 5 and 6) were based on a motion detector that combines preferred-direction enhancement and null-direction suppression, resembling a hybrid of a Hassenstein-Reichardt half-detector and a Barlow-Levick detector, as suggested in [S12].

Stimuli were simulated in a 2-dimensional space covering 90° in both azimuth and elevation with 1° resolution. Each hypothetical motion (half-)detector had three neighboring input lines (termed A, B and C) which were offset by 5° from each other along the horizontal axis (for simplicity). Each input line consisted of a spatial and a temporal filter that was applied to the stimulus before further processing. The spatial filter was modeled as a 2D convolution with a Mexican hat filter kernel using the above definition (see "Spatial receptive field model") and the fitted parameters from table S1 and S2. The temporal filter consisted of either a 1st order band-pass or as a 1st order low-pass filter with the time constants from the table correspondingly. Subsequent rectification simulated the polarity selectivity of the input lines to the downstream motion detector. To implement the nonlinear interaction between the three input lines in the most simplified, we modelled the nonlinear action as $A \cdot B/(C + 0.1)$ involving only one free parameter to avoid division by zero. 270 of these elementary motion detectors were arranged on a 2-dimensional grid, separated by 5° from each other.
To evaluate the performance and tuning of the simulated detectors across stimulus frequencies, we measured the mean response of the simulated (half-)detectors to moving gratings at different speeds. Vertically oriented square wave gratings of 24° wavelength were swept over the detector array with 50 different velocities corresponding to 50 different contrast frequencies

logarithmically spaced between 0.01 Hz and 20 Hz. The gratings moved for 5s to the right followed by a pause of 0.5s and 5s of motion in the opposite direction. The time step for all simulations was 10 ms. The direction of the stimulus that elicited the strongest response across all frequencies was termed the preferred direction (PD) of the respective motion detector. Consequently, the other direction was the null direction (ND).
The direction selectivity of the resulting tuning curve was evaluated by defining a direction selectivity index (DSI)

$$DSI = \frac{\sum PD - \sum ND}{\sum PD + \sum ND}$$

where the sum goes over all frequencies simulated. This definition produces DSI values between 0 and 1, where 1 means perfect, and 0 means no direction selectivity. Secondly, the optimal frequency f_{opt} was defined as the stimulus frequency that elicited the strongest response in PD direction.
The above measure only quantifies the response difference between the two opposing directions of motion along the main axis of the detector. However, it cannot distinguish between detectors that differ in their response properties to intermediate directions of motion. Hence, we additionally assessed the directional tuning specificity of each detector by measuring its response to differently oriented moving gratings. We stimulated the model with square wave gratings of 24° wavelength, rotated by different angles from 0° to 360° in steps of 30°, and measured the mean response of the detector array at the optimal frequency f_{opt}, as determined above. From the corresponding simulated responses, the direction selectivity was quantified as the length of the normalized response vector:

$$L_{dir} = \left| \frac{\sum_{\varphi} \vec{v}(\varphi)}{\sum_{\varphi} |\vec{v}(\varphi)|} \right|$$

where $\vec{v}(\varphi)$ is a vector proportionally scaled with the mean detector response and pointing in the corresponding stimulus direction of motion given by the rotation angle φ of the stimulus. This quantity L_{dir} has been suggested as a robust measure of direction selectivity that includes both relative response magnitude and tuning width of a direction selective neuron [S13].
For the bar plots in Figure 5B&E (right column) the simulations were repeated also for all possible implementations of a two-arm detector whose nonlinear interaction was either modelled as $A \cdot B$ for a classical Hassenstein-Reichardt-(half-)detector or as $B/(C + 0.1)$ for a Barlow-Levick-detector.
All simulations were performed using Python 2.7.

Statistics

Throughout this article, values are reported as mean ± standard error (SEM). In order to quantify the significance of the effect of CDM application on the temporal response characteristics of the medulla cells, we defined three different measures for the impulse responses: a) the time to the first peak $t_{peak}^{(1)}$ is the time between the onset of the impulse response (defined as the time when it has reached 15% of its maximum value) and the time when it has reached its maximum value; b) the time to the second peak $t_{peak}^{(2)}$ is similarly defined as the time between the onset of the impulse response and the peak of the subsequent undershoot or overshoot, which is defined only for the band-pass filters; c) lastly, we defined a peak width w_{peak} as the width of the first

peak at half maximum. We quantified these values for each fly and tested the change between control and CDM condition for significance using a paired t-test.

Supplemental References

S1. 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.

S2. Strother, J. A., Wu, S.-T., Wong, A. M., Nern, A., Rogers, E. M., Le, J. Q., Rubin, G. M., and Reiser, M. B. The emergence of directional selectivity in the visual motion pathway of *Drosophila*. Neuron *in press*.

S3. 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.

S4. 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.

S5. Pologruto, T. A., Sabatini, B. L., and Svoboda, K. (2003). ScanImage: flexible software for operating laser scanning microscopes. BioMed Eng OnLine *2*, 13.

S6. Reiser, M. B., and Dickinson, M. H. (2008). A modular display system for insect behavioral neuroscience. J Neurosci Methods *167*, 127–139.

S7. French, A. S. (1976). Practical nonlinear system analysis by Wiener kernel estimation in the frequency domain. Biol Cybern *24*, 111–119.

S8. Dayan, P., and Abbott, L. F. (2005). Theoretical Neuroscience. MIT Press.

S9. 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.

S10. Borst, A., and Abarbanel, H. D. I. (2007). Relating a calcium indicator signal to the unperturbed calcium concentration time-course. Theor Biol Med Model *4*, 7.

S11. Schnell, B., Weir, P. T., Roth, E., Fairhall, A. L., and Dickinson, M. H. (2014). Cellular mechanisms for integral feedback in visually guided behavior. Proc Natl Acad Sci USA *111*, 5700–5705.

S12. Haag, J., Arenz, A., Serbe, E., Gabbiani, F., and Borst, A. (2016). Complementary mechanisms create direction selectivity in the fly. eLife *5:e17421*.

S13. Mazurek, M., Kager, M., and Van Hooser, S. D. (2014). Robust quantification of orientation selectivity and direction selectivity. Front Neural Circuits *8*, 92.

2.2 VISUAL PROJECTION NEURONS MEDIATING DIRECTED COURTSHIP IN *Drosophila*

SUMMARY This study investigates a class of visual projection neurons, LC10, that is dedicated to guiding directed courtship behavior in male *Drosophila*.

First, we used the the custom-written automated tracking software *MateBook* to analyze courtship behavior of freely moving flies in a small arena. These experiments established that male fruit flies rely heavily on visual cues to perform directed courtship. We then identified the visual projection neuron LC10 as a candidate cell type to relay visual input signals that are relevant for courtship behavior to the central brain. Using a virtual reality set-up, we could show that LC10 neurons are required for tracking of small visual objects: Blocking the output of these neurons significantly impaired the ability of male flies to follow a virtual bead. Next, calcium imaging experiments in LC10 revealed that these neurons have a preference for moving figures of behaviorally relevant sizes and speeds, while large field motion stimuli suppressed the neuronal activity. By using a stochastic motion noise stimulus, we found that these response properties are implemented through a motion-based center surround mechanism. Finally, unilateral optogenetic activation of LC10 neurons elicited ipsilateral turning and wing extensions, a behavior which is characteristic for directed courtship in male *Drosophila*. We concluded that LC10 neurons relay visual information about potential mating partners to the male courtship circuitry and thus visually guide directed courtship behavior in male fruit flies.

This article was published in *Cell* in July 2018 (Ribeiro et al., 2018). It was highlighted in Nojima et al. (2018).

AUTHORS Inês M.A. Ribeiro, **Michael Drews**, Armin Bahl, Christian Machacek, Alexander Borst and Barry J. Dickson

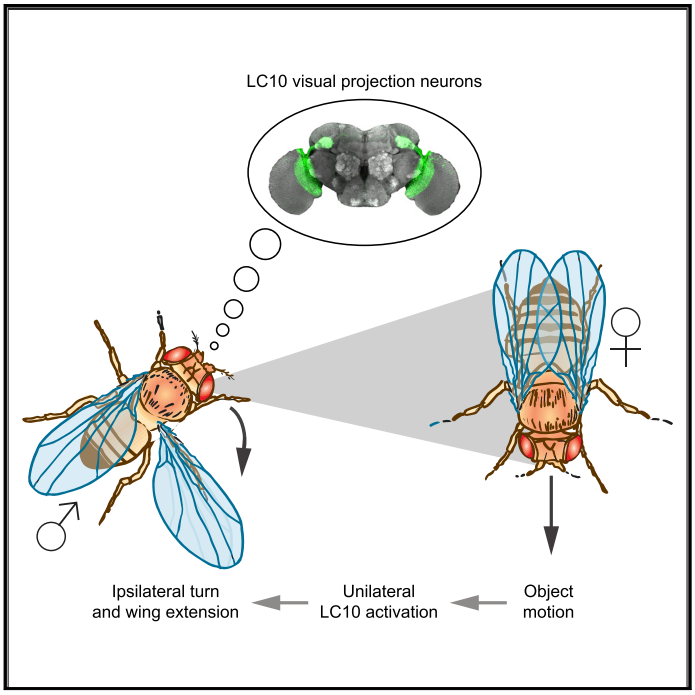
CONTRIBUTIONS I.M.A.R, B.J.D., and A. Borst conceived and designed this study. **M.D.** performed the functional imaging and data analysis presented in Figure 5. I.M.A.R. performed all other experiments and analyzed the data. A. Bahl and **M.D** .conceived and designed the visual stimuli. C.M. and B.J.D. created MateBook. I.M.A.R and B.J.D. co-wrote this paper with input from all authors.

Article

Cell

Visual Projection Neurons Mediating Directed Courtship in *Drosophila*

Graphical Abstract



Authors

Inês M.A. Ribeiro, Michael Drews, Armin Bahl, Christian Machacek, Alexander Borst, Barry J. Dickson

Correspondence

aborst@neuro.mpg.de (A.B.), dicksonb@janelia.hhmi.org (B.J.D.)

In Brief

A specific class of visual projection neurons are dedicated to guiding courtship behaviors in *Drosophila*.

Highlights

- *Drosophila* males rely on LC10 neurons to perform directed courtship
- LC10 neurons are essential for tracking a fly-size visual object but not a long bar
- LC10 neurons sense moving visual objects with a motion-based center surround
- Activation of LC10 elicits directed courtship actions enhanced by courtship arousal

Ribeiro et al., 2018, Cell *174*, 607–621
July 26, 2018
https://doi.org/10.1016/j.cell.2018.06.020

CellPress

Article

Cell

Visual Projection Neurons Mediating Directed Courtship in *Drosophila*

Inês M.A. Ribeiro,[1,2,3] Michael Drews,[2] Armin Bahl,[2] Christian Machacek,[3] Alexander Borst,[2,*] and Barry J. Dickson[1,3,4,5,*]
[1]Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA 20147, USA
[2]Max Plank Institute of Neurobiology, Am Klopferspitz 18, 82152 Martinsried, Germany
[3]Institute of Molecular Pathology, Campus-Vienna-Biocenter 1, 1030 Vienna, Austria
[4]Queensland Brain Institute, University of Queensland, St. Lucia, QLD 4072, Australia
[5]Lead Contact
*Correspondence: aborst@neuro.mpg.de (A.B.), dicksonb@janelia.hhmi.org (B.J.D.)
https://doi.org/10.1016/j.cell.2018.06.020

SUMMARY

Many animals rely on vision to detect, locate, and track moving objects. In *Drosophila* courtship, males primarily use visual cues to orient toward and follow females and to select the ipsilateral wing for courtship song. Here, we show that the LC10 visual projection neurons convey essential visual information during courtship. Males with LC10 neurons silenced are unable to orient toward or maintain proximity to the female and do not predominantly use the ipsilateral wing when singing. LC10 neurons preferentially respond to small moving objects using an antagonistic motion-based center-surround mechanism. Unilateral activation of LC10 neurons recapitulates the orienting and ipsilateral wing extension normally elicited by females, and the potency with which LC10 induces wing extension is enhanced in a state of courtship arousal controlled by male-specific P1 neurons. These data suggest that LC10 is a major pathway relaying visual input to the courtship circuits in the male brain.

INTRODUCTION

Visual detection and localization of moving targets are essential tasks in the life of many animals. For example, a raptor hunting prey (Nikonov et al., 2006), a zebrafish larva capturing paramecia (Budick and O'Malley, 2000), or a housefly chasing a conspecific (Land and Collet, 1974) all use visual cues to adjust their motor output to the movements of their target with sub-second precision. In several systems, the neural circuits involved in detection, location, and tracking of moving targets are beginning to be explored. For example, in large flies, visual neurons projecting to the central brain detect small visual objects and are proposed to aid in object tracking (Gilbert and Strausfeld, 1991; Nordström et al., 2006; Trischler et al., 2007). Similarly, in zebrafish larvae, one relay station receiving inputs from the retina and the optic tectum has been implicated in prey capture (Del Bene et al., 2010; Gahtan et al., 2005). Nonetheless, how vision guides the tracking of a potential mate or prey is still poorly understood. Here, we investigate visual tracking in fruit flies.

Drosophila melanogaster males display an intricate courtship ritual composed of stereotypic actions performed in a variable sequence. These courtship actions include extension and vibration of a single wing to produce courtship song, orienting toward the female, following, tapping, licking, attempted copulation, and copulation (Hall, 1994). Olfactory and gustatory cues signal the species, gender, and receptivity of a potential mate and are thought to influence the male's decision to initiate courtship (Dweck et al., 2015; Kurtovic et al., 2007; Lu et al., 2012; Thistle et al., 2012; Toda et al., 2012). Vision affects courtship success (Krstic et al., 2009; Markow, 1987; Markow and Hanson, 1981; Markow and Manning, 1980) and may also contribute to courtship initiation (Agrawal et al., 2014), but more generally is thought to control the timing and execution of specific actions (Cook, 1979, 1980; Sakai et al., 1997). In particular, vision has been proposed to be essential for the male to effectively chase the female (Cook, 1979, 1980) and to extend the wing closest to the female when singing (Kohatsu and Yamamoto, 2015; Pan et al., 2012); that is, to direct his courtship displays toward the target. Visual cues may also be used to assess the distance to the female, and to accordingly modulate the amplitude of song pulses (Coen et al., 2016).

Visual input consists of a two-dimensional array of light intensities varying across space and time. Several computational steps are needed to extract visual features such as wide-field motion, small visual objects, and looming objects (Borst, 2014). In vertebrates, visual features are relayed to the central brain by retinal ganglion cells, which connect the plexiform layers in the retina with several regions in the central brain (Dhande et al., 2015). Several classes of retinal ganglion cells exist, each presenting a different response profile to such visual features (Baden et al., 2016). These include direction-selective and non-selective cells that sense either local or global motion (Baden et al., 2016; Dhande et al., 2015). The invertebrate analogs of the vertebrate retinal ganglion cells are the visual projection neurons, which connect the neuropils of the optic lobe to the central brain. Many visual projection neurons have been characterized in larger flies. For example, male hoverflies, blowflies, fleshflies, houseflies, and hawk moths pursue their mates in flight (Land and Collet, 1974; Olberg et al., 2000), and in each of these species visual projection neurons have been identified that

detect small targets with angular sizes and speeds matching those of conspecifics (Collet, 1972; Gilbert and Strausfeld, 1991; Nordström et al., 2006; Trischler et al., 2007). Many of these visual projection neurons are sexually dimorphic and receive visual input from the acute zone, a specialized area in the male retina with high visual acuity. Functional characterization of visual neurons during behavior is difficult in these systems, however, and if and how these neurons convey visual cues during mate pursuit remains an open question.

Recently, studies in *Drosophila* have begun to define the many distinct classes of visual projection neurons and link their detection of specific visual features with behavioral output (Haikala et al., 2013; Kim et al., 2017; von Reyn et al., 2017; Wu et al., 2016). Many classes of visual projection neurons have been described based on the optic lobe neuropils they innervate and the optic glomeruli they target in the lateral protocerebrum of the central brain (Otsuna and Ito, 2006; Wu et al., 2016). Large tangential cells projecting from the lobula plate sense large field optic flow and are thought to regulate flight stabilization (Joesch et al., 2008; Kim et al., 2017; Maimon et al., 2010). Lobula columnar (LC) neurons of various subtypes tile the lobula neuropil (Otsuna and Ito, 2006; Wu et al., 2016), some of which have been assigned functional roles. LC6 and LC16 neurons sense looming stimuli and when artificially activated elicit escape, by jumping or backward walking respectively (Wu et al., 2016). LC11 neurons specifically sense small dark objects (smaller than 10°) (Keleş and Frye, 2017) but have not yet been shown to function in target tracking. Indeed, unlike larger flies, fruit flies court on food substrates at short distances from the target, at which conspecifics are more likely to subtend large angular sizes (above 20°) on the male retina, outside the size range sensed by LC11.

We hypothesized that courtship-relevant visual information might be conveyed by a different subset of visual projection neurons and set out to identify and characterize these neurons. We present evidence that the LC10 class of visual projection neurons responds to moving targets that have angular sizes and speeds matching those of the female during courtship, and that their activity is both necessary and sufficient for visually guided aspects of male courtship behavior. Our work clarifies the contributions of visual input to male courtship behavior in *Drosophila* and provides a causal link between a major visual pathway and directed courtship.

RESULTS

Automated Video Analysis Defines the Role of Motion Vision in Directed Courtship

We developed the MateBook software to automate the analysis of courtship videos (Figure S1; STAR Methods). MateBook uses machine vision techniques to track the trajectories and wing positions of each of two flies in circular chambers and derive a number of fly-specific and pairwise attributes for each video frame (Table S1). Specific courtship actions are detected by applying user-defined classifiers. A challenge in the automated analysis of courtship is the high frequency of occlusions, in which the initial image segmentation yields only a single merged object. This makes it difficult to assign wings to individual flies during occlusions, and to maintain individual fly identities across occlusions. MateBook only reports wing positions during those occlusions in which the merged object can reliably be resolved into two flies. To assign fly identities across occlusions, MateBook relies on the distinct sizes of the two flies (females are typically larger than males) and an estimate of the more likely trajectories through the occlusion. Its accuracy in resolving occlusions depends on the assay and video format. For all the assays reported here, we used chambers of 18 mm diameter and acquired video recordings at 25 Hz. In a test set of 15 manually annotated videos of wild-type males courting wild-type females, 3,140 (96%) of 3,275 occlusions were correctly resolved. Because identity assignment relies on size differences that are detectable in non-occluded frames, misassignments are generally confined to shorter sequences between two occlusions and not propagated through the entire video. Importantly, the courtship statistics reported by MateBook for this test set of videos were indistinguishable between the datasets generated using automated or manual identity assignments (Figures 1A, 1B, 1F, S2A, and S2H). Accordingly, we relied exclusively on fully automated video analysis by MateBook for all data reported here.

We used MateBook to reexamine the role of vision in courtship behavior, assessing males that were completely blind (*ort* mutants, which lack the ionotropic histamine-gated chloride channel required in lamina monopolar cells for neurotransmission from all photoreceptor cells R1-R8) (Bulthoff, 1982), lack visual input from photoreceptors R1 to R6 (*ninaE* mutants, which lack Rhodopsin1 (Rh1) in photoreceptors R1–R6) (Scavarda et al., 1983), or motion-blind due to expression of tetanus toxin light chain (TNT) (Sweeney et al., 1995) in the elementary motion-detecting neurons T4 and T5 ("T4T5 block" flies) (Schnell et al., 2012). The most pronounced courtship deficit in each of these visually impaired males was a dramatic reduction in their ability to orient toward and remain close to the female (Figures 1C–1E, 1G–1I, S2B, and S2C), and hence less frequent detection of following, courting, and copulation events (Figures S1C–S1E and S1H–S1J). For wild-type control males, the distributions of the angle toward and the distance from the female were highly biased to small values (but not zero, due to the curvature of the chambers and the fly body length, respectively; Figure 1B). In contrast, for the completely blind *ort* males, no such bias was observed and the distributions of the angles and distances to the female were indistinguishable from those obtained from data in which the female positions were shuffled (Figure 1C). R1–R6-blind *ninaE* males were only marginally better at orienting and following than *ort* males (Figure 1D); T4T5-block males did show some bias for low angles and distances, yet still well below the level observed with control males (Figures 1E and S2B).

Despite the severe reduction in orienting and following by each of these visually impaired males, the frequency of wing extension was only mildly affected, if at all (Figure S1I). Males preferentially sing by vibrating the wing that is oriented toward the female (Kohatsu and Yamamoto, 2015; Pan et al., 2012), as we confirmed in our analysis of wild-type males by plotting the distribution of the female's position when a wing is extended (Figure 1F), and by computing a wing choice index ([ipsilateral – contralateral wing extensions]/total wing extensions; Figure 1F). This ipsilateral bias is lost in *ort* and *ninaE* mutants and T4T5-block flies (Figures 1G–1I). We conclude that vision, and in particular

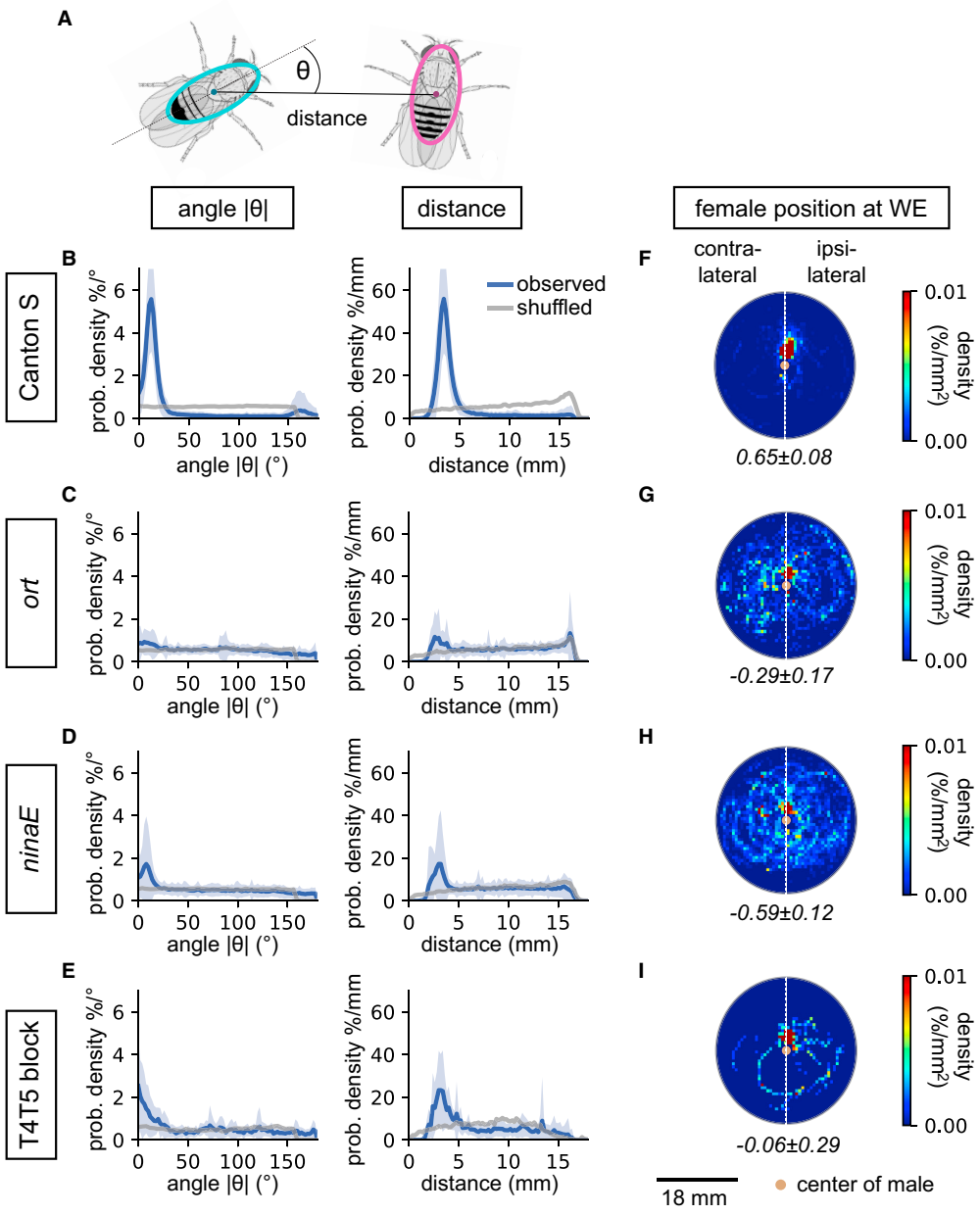


Figure 1. Visual Input Mediates Directed Courtship

(A) Angle θ and distance from male to female.

(B–E) Probability density functions of the absolute angle θ and distance for single-pair courtship assay between males of the indicated genotype and *Canton S* females, for real (blue) and shuffled (gray) data. (B) *Canton S* males, n = 93. (C) ort^{1} males, n = 57. (D) $ninaE^{P332}$ males, n = 108. (E) T4T5 block (*SS324>TNT*) males, n = 23.

(F–I) Heatmap for the position of the female relative to the male for every frame in which the male extends a wing. The wing choice index is indicated below the heatmap (STAR Methods; mean ± SEM). (F) *Canton S* males, n = 93. (G) ort^{1} males, n = 57. (H) $ninaE^{P332}$ males, n = 108. (I) T4T5 block (*SS324>TNT*) males, n = 23. See also Figures S1 and S2 and Tables S1 and S2.

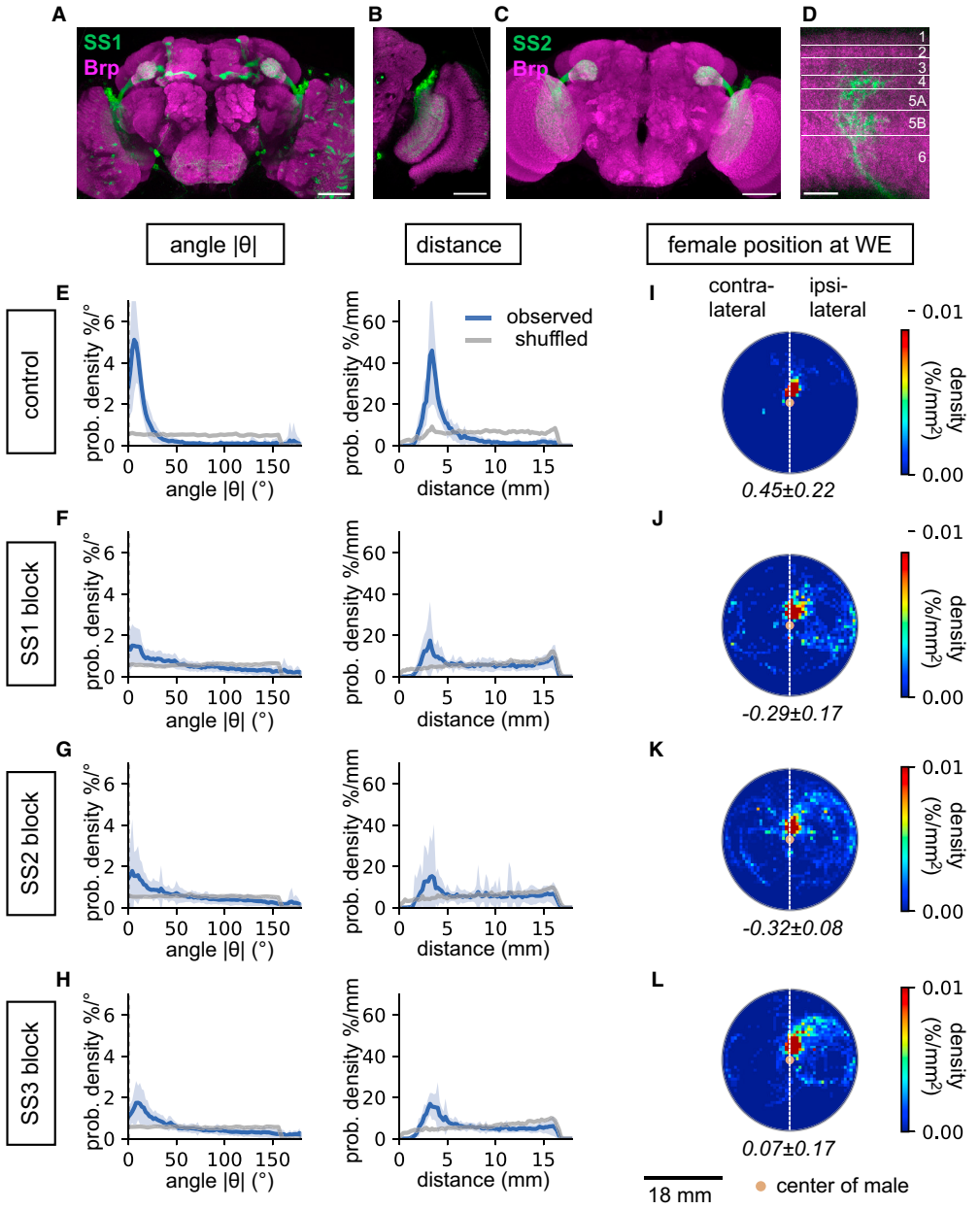


Figure 2. LC10 Visual Projection Neurons Are Required for Directed Courtship

(A–C) Expression pattern of *LC10-SS1* (A and B) and *LC10-SS2* (C) visualized with mCD8-Tomato reporter and anti-RFP (green), counterstained with anti-Brp (magenta) to label all synapses. (A and C) Anterior view of the brain. (B) Slice of the optic lobe at the level of the lobula. There are 128.4 ± 2.5 (mean ± SEM) LC10 cells, n = 6 for the *LC10-SS1* driver. Scale bar, 50 μm.

(D) Lobula arborizations of a single cell in *LC10-SS1* labeled using the multi-color flip out (Nern et al., 2015) and visualized with anti-FLAG (green) and anti-Brp (magenta) antibodies. Scale bar, 10 μm.

(E–H) Probability density functions of the absolute angle θ and distance for single-pair courtship assay between males of the indicated genotypes and *Canton S* females for real (blue) and shuffled (gray) data. (E) *LC10-SS1* > *IMPTNT-Q* (inactive TNT); n = 19. (F) *LC10-SS1* > *TNT*; n = 27. (G) *LC10-SS2* > *TNT*; n = 100. (H) *LC10-SS3* > *TNT*; n = 30.

(legend continued on next page)

Rh1-dependent vision, is essential for males to direct their locomotion and wing extensions toward the female during courtship.

Visual Projection Neurons Required for Directed Courtship

We searched for visual projection neurons that might relay courtship-relevant visual features from the optic lobes to the central brain. Reasoning that such neurons, like many involved in courtship behavior, might express the *fruitless* (*fru*) gene (Kimura et al., 2008), we initially focused on fru^+ visual projection neurons by screening for courtship deficits in flies in which TNT was expressed in cells co-labeled by *fru-FLP* and various VT enhancer GAL4 lines that were known to target visual projection neurons (Tirian and Dickson, 2017). Four positive lines were identified (*VT043656*, *VT040012*, *VT047880*, *VT002224*), each of which, when combined with *fru-FLP* labeled 15–30 neurons that connect the lobula and the anterior optic tubercle (AOTu) (Figure S3A). These cells appeared to be a subset of the neuronal class that was previously described as Lv1+Ld (Kimura et al., 2008) or LC10 (Aptekar et al., 2015; Otsuna and Ito, 2006; Wu et al., 2016) cells; we use the latter designation here. We used the split-GAL4 technique (Luan et al., 2006; Pfeiffer et al., 2010) to derive three independent driver lines specific for the LC10 cells, which we refer to as *LC10-SS1*, *LC10-SS2*, and *LC10-SS3* (Figures 2A–2C and S3B; Table S2). They each label ~100 cells per hemisphere.

A strong reduction in directed courtship was observed when we expressed TNT with any of these three drivers (Figures 2E–2L, S1F, and S1G). These LC10-silenced males were impaired in their ability to orient toward and remain close to the female during courtship (Figures 2E–2H and S2C–S2E). When they sang, they did so without any bias for the ipsilateral wing (Figures 2I–2L and S2J–S2L). These deficits are most likely due to a difficulty in locating the female, not a reduced motivation to court, since males with silenced LC10 neurons vigorously courted females immobilized by decapitation (Figures S3E–S3H).

Four LC10 subtypes have been defined, based on their distinctive dendritic arborizations in the lobula (Wu et al., 2016). *LC10-SS1* mostly labels the LC10a subtype, which arborizes in lobula layers 3 to 5 with lateral extensions in layers 4 and 5B (>30 cells for SS1, Figures 2D and S3C). Most cells labeled by *LC10-SS2* and *LC10-SS3* also appear to belong to the LC10a subtype, with some additional cells with arborizations spanning larger areas of the lobula, likely of LC10b subtype (data not shown) (Wu et al., 2016). We also found that ~60% of the cells labeled in *LC10-SS1* co-express *fru-LexA* (Figure S3D) (Mellert et al., 2010), suggesting that our initial intersections with *fru-FLP* likely targeted these fru^+ LC10a neurons. The common silencing phenotype observed using each of these intersectional strategies thus most likely reflects the function of fru^+ LC10a neurons, although we cannot exclude an additional contribution from either fru^- LC10a neurons or LC10b neurons (Figure S4). We could not detect any sex differences in the morphology of LC10 neurons. To test for possible functional differences, we genetically feminized them using *LC10-SS1* or *LC10-SS2* to drive a *UAS-tra* transgene (Billeter et al., 2006; Boggs et al., 1987; Ferveur et al., 1995). Males with feminized LC10 neurons tracked and courted females just as well as control males did (Figures S3I–S3Q).

LC10 Neurons Are Required for Visual Object Tracking

Neural processing within the optic lobe is thought to extract relevant visual features from the spatial array of retinal inputs, which are delivered to the central brain by visual projection neurons. To determine which visual features might be relayed by LC10 neurons, we probed the responses of LC10-silenced flies to a range of specific visual stimuli. Tethered males were placed on an air-suspended ball, gently warmed to instigate robust walking, and presented with visual stimuli displayed on three computer screens placed around the fly (Bahl et al., 2013; Seelig et al., 2010). In this set up, LC10-silenced males were indistinguishable from controls in their ability to turn in the same direction as a rotating full-field grating (Figures 3A, 3B, S5A, and S5B) or a long lean bar sweeping the screen in open loop (Figures 3C–3F and S5C–S5F). In a closed-loop setting, they also fixated the bar just as well as control males (Figures S5G and S5H). LC10 neurons are thus dispensable for the optomotor response and for long bar fixation.

Small bars did not elicit a strong turning response in open loop, possibly due to the uncoupling of visual and proprioceptive information in the open-loop configuration (Fujiwara et al., 2017). We therefore tested responses to smaller objects in a closed-loop configuration. We developed a three-dimensional closed loop stimulus consisting of a virtual cylinder in which a rectangular fly-sized virtual object moves around a circle of 25 mm diameter at a constant speed of 7.85 mm/s, within the speed range of females in a courtship arena (Figure 3G). The position, size, and speed of this virtual object were adjusted according to the male's locomotion within the virtual cylinder. Wild-type and control males tracked a dark virtual object, keeping it centered and in close virtual distance (Figures 3H, 3I, and S5I–S5L). Curiously, control males also tracked a bright virtual object but aimed for its edges, resulting in high probabilities of the bright virtual object being kept close but both central and lateral to the male (Figures 3L, 3M, and S5M–S5P). In contrast, LC10-block males did not track either dark or bright virtual moving objects (Figures 3J, 3K, 3N, and 3O). The overall locomotion of LC10-block males was indistinguishable from control males when presented with an invisible virtual object (Figures 3P–3S and S5Q–S5T), as were the distributions of translational and rotational speeds across tests (Figures S5U and S5V). We conclude that LC10 neurons are specifically required to track the virtual object.

Functional Imaging of LC10 Neurons Reveals a Preference for Visual Figures

A moving virtual object presents a number of different visual features, such as its motion, angular size, or angular position. We

(I–L) Heatmap of the position of the female relative to the male for every frame in which the male extends a wing. The wing choice index is indicated below the heatmap (STAR Methods; mean ± SEM). (I) *LC10-SS1* > *IMPTNT-Q* (inactive TNT); n = 19. (J) *LC10-SS1* > *TNT*; n = 27. (K) *LC10-SS2* > *TNT*; n = 100. (L) *LC10-SS3* > *TNT*; n = 30.

See also Figures S1, S2, S3, and S4 and Tables S1 and S2.

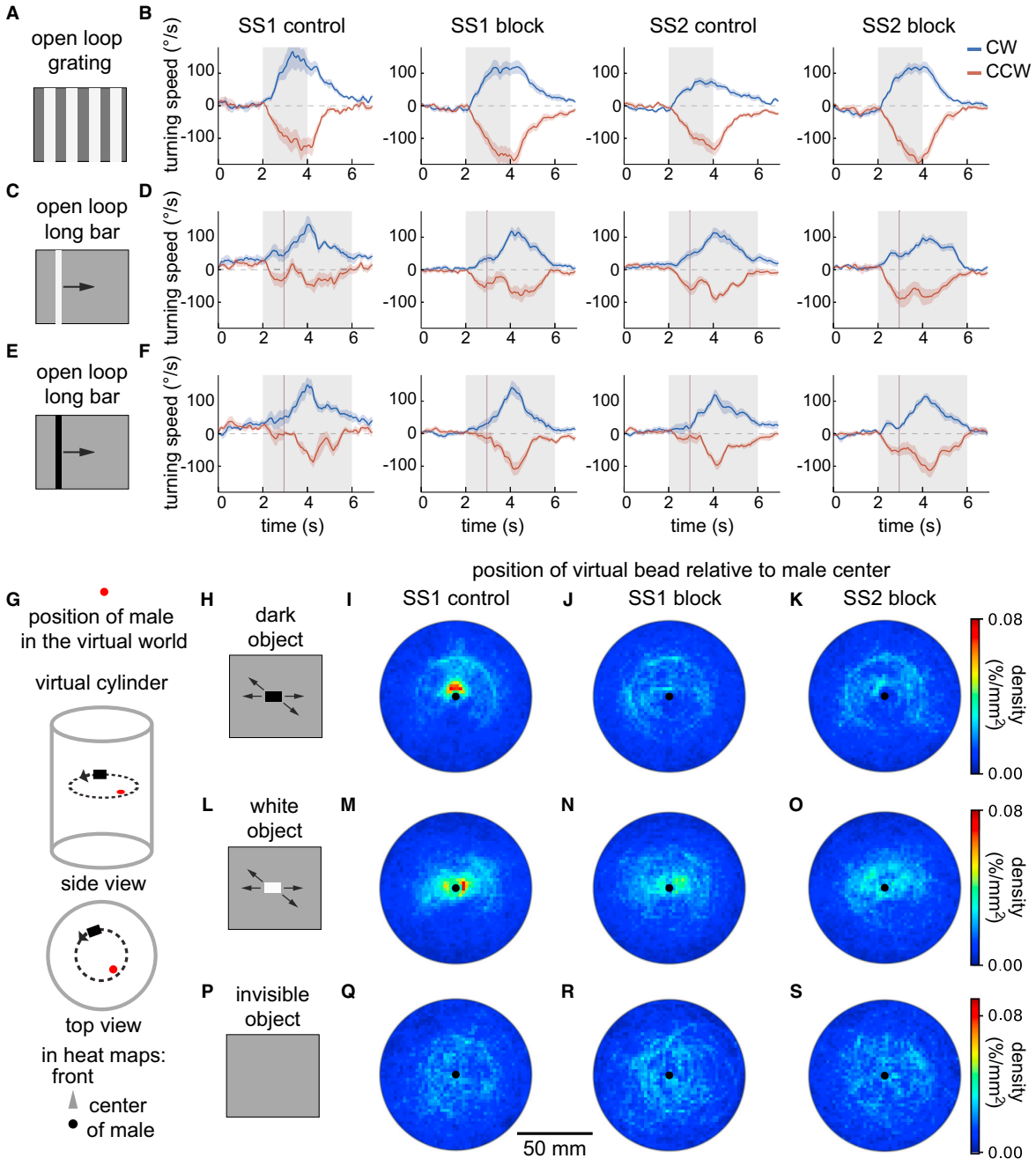


Figure 3. LC10 Neurons Are Required for Specific Visual Behaviors

(A, C, and E) Visual stimuli presented in open loop to the tethered male walking on an air-suspended ball. (A) Vertical gratings with wavelength of 20° at 20°/s. (C and E) Vertical 10° bright (C) or dark (E) bar at 60°/s.

(B, D, and F) Turning speed over time in response to visual stimuli shown in (A), (C), (E), respectively. Data are mean ± SEM responses to clockwise (blue) or counter-clockwise (red) motion; gray shade indicates time of motion. SS1 control: *LC10-SS1* > *IMPTNT-Q*, n = 5; SS1 block: *LC10-SS1* > *TNT*, n = 9; SS2 control: *LC10-SS2* > *IMPTNT-Q*, n = 11; SS2 block: *LC10-SS2* > *TNT*, n = 7.

(legend continued on next page)

performed functional imaging to determine the visual features that activate LC10 neurons. Immobilized males expressing the calcium indicator GCaMP6m under control of the *LC10-SS1* line were presented with visual stimuli while their calcium responses were recorded with a 2-photon microscope (Chen et al., 2013; Maisak et al., 2013). To avoid mixing calcium responses from multiple cells in the same trace, we selected regions of interest that included only a single punctum within either the lobula or the AOTu (Figures 4A and 4B).

LC10 neurons responded more strongly to moving visual figures, such as a square or long bar (Figures 4C–4F), than to full-field stimuli, such as gratings that moved vertically or horizontally (Figures 4G and 4H) (Aptekar et al., 2015). LC10 calcium responses were similarly strong for square bars of either contrast, but markedly stronger for bright long bars than for dark long bars. For the visual figures, the direction of movement had only a modest impact on the calcium response (Figure 4I). No response was observed with flickering gratings (Figure 4J), but, surprisingly, OFF full-field flicker elicited a detectable but slow increase in the calcium signal in LC10 neurons, most notably in their dendrites (Figure 4K). Both bright and dark looming discs elicited a calcium response in LC10 neurons, with rapidly looming stimuli eliciting much faster transients than slowly looming stimuli (Figures 4L–4O). Responses of LC10 neurons in females were qualitatively and quantitatively similar to their responses in males (Figures S6D–S6O). In summary, LC10 neurons respond more strongly to various moving objects than they do to full-field motion stimuli.

A Motion-Based Center-Surround Mechanism for Figure Detection by LC10 Neurons

LC10 terminals are retinotopically arranged in the AOTu (Wu et al., 2016), suggesting that LC10 neurons might provide positional information of moving objects. If so, we would expect that LC10 dendrites collect visual information from a limited area within the visual field. To determine the LC10 receptive field, we presented random white noise patterns of 5° by 5° squares covering the entire stimulus arena (Figure 5A). This motionless white noise stimulus led to small calcium transients in the LC10 arborizations in the lobula, indicating that LC10 neurons respond to local luminosity changes. Cross-correlation between the time series of random checkered patterns and LC10 calcium responses yielded a local OFF-selective receptive field spanning an area of ~25° (Figure 5B).

Because the calcium transients evoked by local flicker were much smaller than those elicited by moving bars, we reasoned that other local visual features might contribute to figure detection by LC10 neurons. Naturalistic visual scenes as well as artificial visual stimuli contain many local edges—the sharp, moving borders between two areas of the visual scene with opposing contrasts. By presenting single full-field moving edges with opposing contrasts, we determined that LC10 neurons are indeed sensitive to edges. These stimuli triggered sharp but low calcium transients in the LC10 dendrites and axon terminals (Figure S6A), without a clear preference for direction (Figures S6B and S6C). In LC10 dendrites, dark edges elicited a low and slowly decaying calcium transient whereas bright edges elicited a strong and rapidly decaying response. Importantly, calcium transients evoked by moving edges were stronger than those evoked by local flicker, suggesting that edge motion constitutes an important input to LC10 neurons.

How might LC10 neurons detect small objects? A plausible hypothesis is that LC10 neurons might be part of a system that discriminates moving objects from the background by sensing local motion cues that differ from the global motion of the visual scene. The hallmark of such mechanism would be a receptive field selective to motion at one spot but inhibited by motion around that spot, i.e., a motion-based antagonistic center-surround configuration of the receptive field (Reichardt and Poggio, 1976). To test this, we presented a motion noise stimulus consisting of 10 sine gratings with a wavelength of 20°, moving independently within 10 stripes at different elevations in the visual arena (Figure 5C). Each grating followed a random velocity profile with a mode of 20°/s (Figures 5D and 5E), creating a stimulus with stochastic velocity within each stripe and uncorrelated motion between stripes (Figure 5F). Calcium transients recorded in isolated puncta in LC10 arborizations were positively correlated with the speed at the center of the receptive field at a specific elevation (red in Figures 5G, S6P, and S6Q), with a milder inverse correlation in neighboring elevations (blue in Figures 5G and S6Q). The observed temporal component of the receptive field was slower than the stimulus autocorrelation, indicating that the speed of the motion noise stimulus did not distort our estimation of the temporal component (Figures 5H, S6R, and S6U–S6Z). The spatial component of this motion-based receptive field presents a local antagonistic center-surround configuration, with a high correlation to speed in the center and a lower and inverse correlation in the surround (Figures 5I, S6S, and S6T). We conclude that LC10 neurons are preferentially sensitive to visual figures due to a motion-based center surround component in their response. The observed low calcium responses to long dark bars are in agreement with such a mechanism. The strong calcium responses to long bright bars are not, possibly indicating that the motion-based center surround observed in LC10 neurons relies on motion in the OFF channel.

LC10 Neurons Are Tuned to Visual Figures with Behaviorally Relevant Sizes and Speeds

We calculated the projection of the ellipse that MateBook fitted to the female onto the male retina for every frame across a total of 6 min of courtship between wild-type males and females. The computed azimuthal angular size of the female ranged from 10° to 140°, with a peak at around 80° that was clearly distinct from shuffled data (Figures 6A and 6B) and reflects the male's tendency to maintain close proximity to the female. Angular speeds were consistently lower than 70°/s, with a slightly wider range in

(G) Virtual cylinder arena and virtual moving object.
(H, L, and P) The three-dimensional virtual object was either dark (H), bright (L), or invisible (P).
(I–S) Heatmaps of probability density function for the position of the dark (I–K), bright (M–O), or invisible (Q–S) object relative to the male center. (I, M, and Q) *SS1>IMPTNT-Q*; n = 9. (J, N, and R) *SS1>TNT*; n = 7. (K, O, and S) *SS2>TNT*; n = 7.
See also Figure S5 and Table S2.

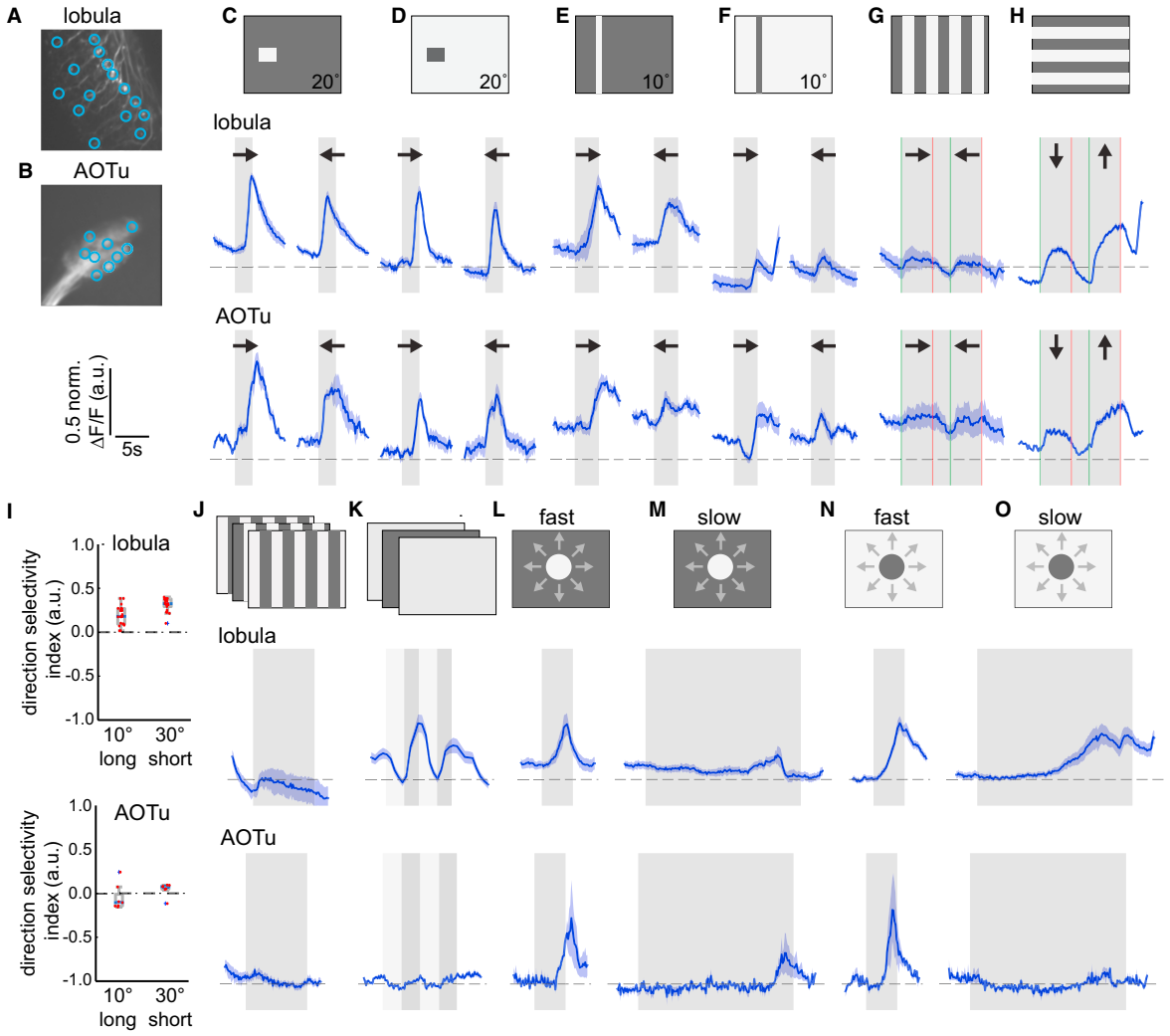


Figure 4. LC10 Neurons Are Sensitive to Small, Moving Visual Objects

(A and B) Single two-photon confocal planes in the lobula (A) and AOTu (B) of an *LC10-SS1>GCaMP6m* male. Circles indicate individual regions of interest (ROIs) selected for quantification of calcium responses.

(C–H) Visual stimuli (top) and calcium responses (shown as mean ± SEM) in LC10 processes in the lobula (middle) and AOTu (bottom). Bright (C) or dark (D) 20° square with progressive or regressive motion with absolute velocity 70°/s. (C) Lobula: n = 122 ROIs in 12 flies; AOTu: 10 ROIs in 2 flies. (D) Lobula: n = 39 ROIs in 4 flies; AOTu: 17 ROIs in 3 flies. Bright (E) or dark (F) 10° bar moving at 50°/s. Vertical (G) or horizontal (H) gratings with wavelength of 20° moving at 20°/s. Data shown in (E)–(H) were acquired from the same samples: lobula: n = 46 ROIs in 4 flies; AOTu: 31 ROIs in 5 flies.

(I) Direction selectivity index (STAR Methods) shown for 30° square bright bar and 10° long bright bar; boxplot shows the median, upper, and lower quartiles.

(J) Counterphase flicker with wavelength of 20° and apparent motion of 20°/s. Data acquired on the same samples as (E)–(H).

(K) Full-field flicker in which fully bright (ON) alternated with fully dark (OFF) visual arena every 3 s; lobula: n = 29 ROIs in 5 flies; AOTu: n = 12 in 4 flies.

(L–O) Bright (L and M) or dark (N and O) looming spheres centered on the receptive field of the cells, with initial angular size of 10° and expanding to full arena at a speed of either 10°/s (L and N) or 2°/s (M and O). Lobula: n = 119 ROIs in 10 flies; AOTu: 20 ROIs in 3 flies.

See also Figure S6 and Table S2.

the observed data compared to shuffled data (Figures 6A and 6B). These tendencies to larger angular sizes and a wider range of angular speeds were even more pronounced in frames classified as courtship (Figure 6C). To assess whether LC10 neurons might be tuned to the visual cues presented by a female during courtship, we examined their calcium responses to short bright bars ranging in angular size from 10° to 80°, moving at absolute angular speeds of up to 550°/s. We found that LC10 neurons are

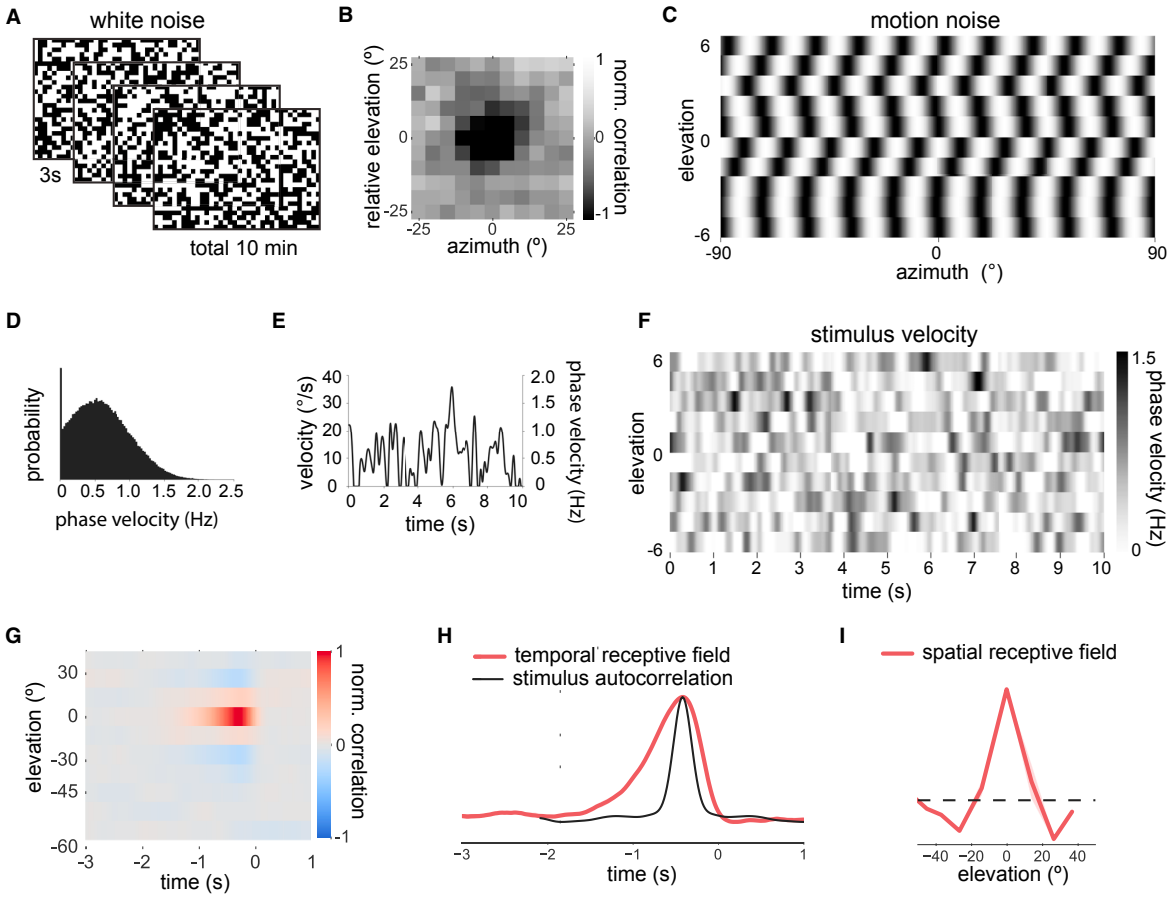


Figure 5. LC10 Neurons Present a Motion-Based Center Surround

(A) White noise stimulus with random checkered patterns of bright and dark 5° by 5° squares covering the entire visual arena changed every 3 s.

(B) Correlation of calcium transients in the LC10 lobula arborizations (*LC10-SS1>GCaMP6m*) with brightness levels of white noise stimulus (n = 4 flies).

(C–F) Statistical characteristics of the motion noise stimulus. (C) Snapshot of the visual arena. (D) Probability distribution of phase velocity. (E) Temporal profile of velocity and phase velocity for one elevation. (F) Velocity profile over a 10-s period for each elevation.

(G) Correlation of calcium transients observed in LC10 arborizations with spatiotemporal velocity profiles. The mean of aligned spatiotemporal profiles is shown, n = 28 ROIs in 5 flies.

(H) The temporal receptive field and the peak-aligned stimulus autocorrelation as a function of time.

(I) The spatial receptive field as a function of relative elevation.

See also Figure S6 and Table S2.

broadly tuned with respect to size but more narrowly tuned with respect to speed, with a preference for stimuli moving in either direction at 150°/s or less (Figures 6D–6F). LC10 neurons are thus broadly tuned to moving figures across a range of angular sizes and speeds that match those a female is likely to present during courtship behavior.

Activation of LC10 Neurons Elicits Directed Courtship Actions that Are Enhanced in a P1-Induced State of Arousal

The data presented thus far indicate that LC10 neurons are likely to be intermittently activated during courtship, depending on the location and movement of the female, and their activity is necessary for the male to turn toward and follow the female and to preferentially use the ipsilateral wing to sing. If LC10 neurons represent the relative position of the female, then unilateral activation of LC10 neurons in an isolated male should be sufficient to trigger ipsilateral turning and wing extension. We tested this by stochastically expressing the optogenetic activator CsChrimson in subsets of LC10 neurons using the *LC10-SS1* driver (STAR Methods) (Klapoetke et al., 2014; Wu et al., 2016). Individual males were placed in a courtship arena illuminated with weak blue light to monitor any behavioral response and exposed to a 0.5-s, 2-s, or 10-s pulse of 627 nm light to activate CsChrimson.

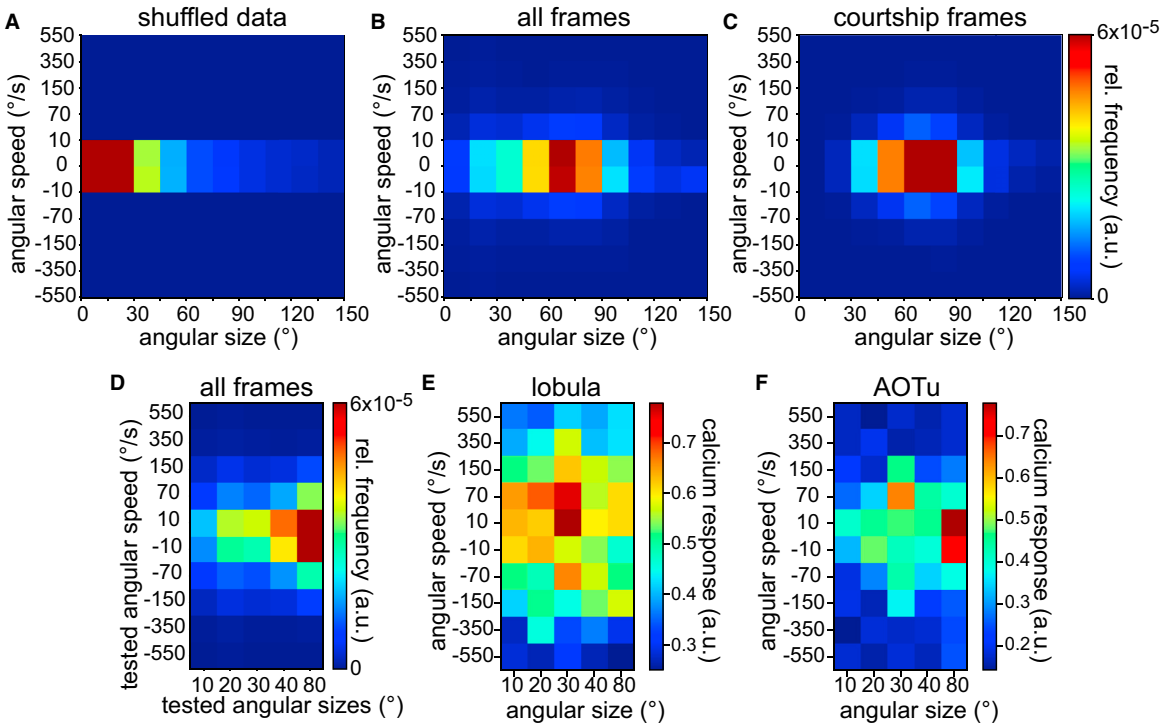


Figure 6. LC10 Neurons Are Sensitive to Objects with Naturalistic Angular Sizes and Speeds

(A–D) Heatmap of distribution of azimuthal angular sizes and speeds the female subtends on the male retina during single-pair courtship assays (n = 93 *Canton S* males paired with *Canton S* females). (A) Shuffled male-female pairs. (B–D) Unshuffled data for all frames excluding copulation (B and D), or all frames in which MateBook detected orienting, following, or wing extension (C). In (D), data are rebinned according to the angular sizes and speeds tested in functional imaging (E and F) such that each bin contains the tested angular size or speed ±10°.

(E and F) Average maximum normalized calcium responses in LC10 processes in the lobula (E, n = 46 ROIs in 6 flies) and AOTu (F, n = 37 ROIs in 4 flies) in *LC10-SS1>GCaMP6m* males to moving bright bars of various speeds and sizes.

See also Table S2.

Subsequently, their brains were dissected to determine post hoc whether CsChrimson was expressed in LC10 neurons, and if so, on which side (Figures S7A and S7B). Males in which LC10 neurons were found to have been unilaterally activated turned ipsilaterally without any noticeable increase in forward velocity (Figures 7A–7D and S7C). A similar unilateral turn was observed in males in which CsChrimson was expressed bilaterally but activated unilaterally by focal stimulation on the ball set up (Figures S7D–S7F). The longer 10-s light pulses also elicited brief wing extensions (lasting 120–240 ms), which were, as predicted, strongly biased to the ipsilateral side (Figure 7E). Thus, in isolated males, unilateral activation of LC10 neurons elicits the same motor responses—ipsilateral turning and wing extension—that are lost when LC10-silenced males court females.

LC10 neurons are functionally monomorphic yet contribute to a male-specific behavior, suggesting that visual information conveyed through LC10 is processed in a sex-specific manner in the brain. A set of brain neurons that might influence such dimorphic processing are the male-specific P1 cells (Kimura et al., 2008), the activity of which is both necessary and sufficient for robust courtship behavior (Kohatsu et al., 2011; Pan et al., 2012; von Philipsborn et al., 2011). Notably, acute activation of P1 neurons elicits courtship that persists for many minutes after the activating stimulus, suggesting that P1 activation induces a lasting state of courtship arousal (Bath et al., 2014; Inagaki et al., 2014). We hypothesized that the ability of LC10 activation to elicit courtship actions might be modulated by this P1-induced state of arousal.

We identified an *LC10-LexA* line that drives expression almost exclusively in LC10 neurons (Figures S7G and S7H), allowing us to test this prediction using a *P1-GAL4* driver (von Philipsborn et al., 2011) to induce the arousal state. Bilateral optogenetic activation of all LC10 neurons using either the *LC10-SS2 GAL4* driver or the *LC10-LexA* driver induced single wing extensions at similar frequencies (Figure 7G), slightly below the frequency with which they had occurred upon unilateral activation (Figure 7E). We did not observe any turning in these experiments, presumably due to the bilateral activation, nor the "reaching" phenotype that has previously been reported to result from LC10 activation (Wu et al., 2016; reaching was however

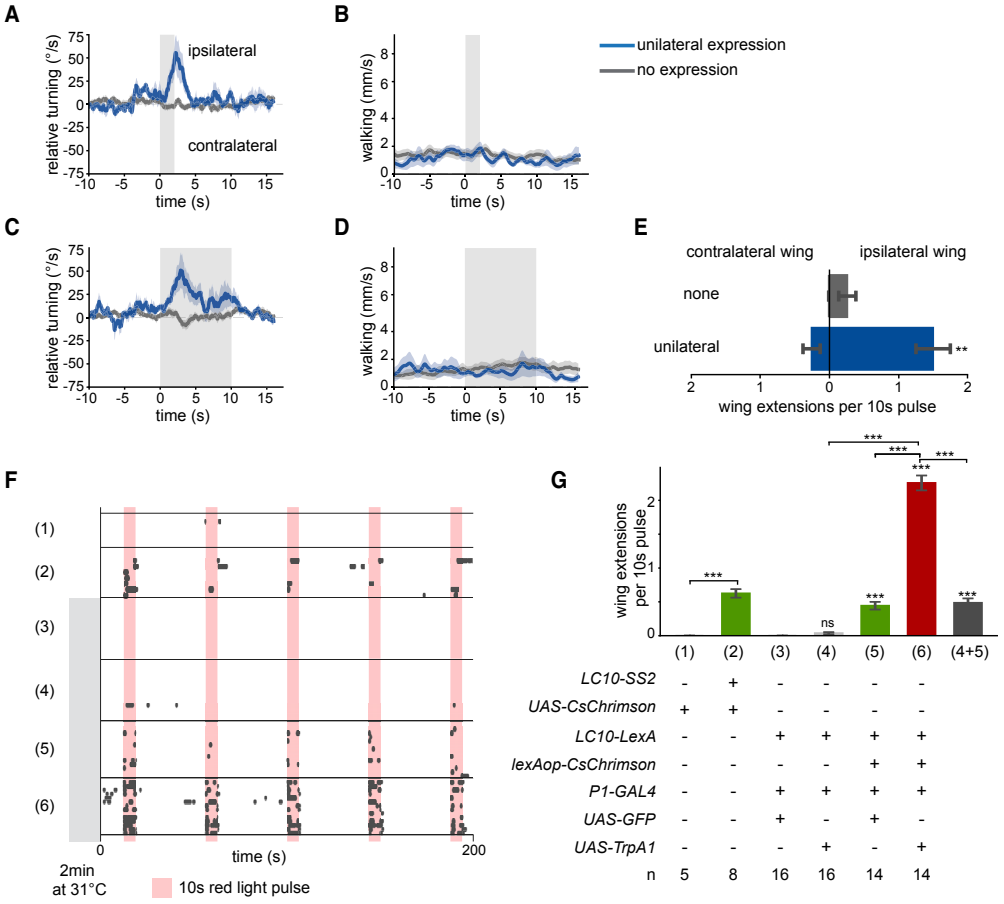


Figure 7. LC10 Activation Elicits Ipsilateral Turning and Short Wing Extensions that Are Potentiated by Pre-activation of P1 Neurons

(A–D) Walking and relative turning speeds of males with unilateral expression (n = 4) or no expression (n = 19) of CsChrimson in LC10 neurons upon 2 s (A and B) or 10 s (C and D) of LED stimulation (*LC10-SS1>FRT-STOP-FRT-CsChrimson*). Positive turning speed means turning to the side of stimulation, shown as mean ± SEM.

(E) Number of short wing extensions elicited by 10 s pulses in males with unilateral (n = 4) or no expression (n = 19) of CsChrimson in LC10 neurons of males shown in (A-D). **p < *0.01*, ns p > *0.05*, Student's t test with Bonferroni correction.

(F) Raster plots of wing extensions per male. Red shades indicate 10-s pulses of stimulation with red LEDs. Grey shade indicates the end of 2 min incubation at 31°C in a dark chamber. Wing extensions were scored immediately thereafter. Vertical line at t = 0.

(G) The average number of wing extensions per 10-s pulse of LED light for the data represented in F, shown as mean ± SEM. ***p < 0.001, ns p > 0.05, Student's t test with Bonferroni correction.

See also Figure S7 and Table S2.

observed when males were tested in groups, as in this previous report). These males are not expected to be in a state of courtship arousal. To induce this state, we pre-activated P1 cells thermogenetically using TrpA1 by heating flies to 31°C for 2 min, just prior to the series of light pulses used for optogenetic activation of LC10 neurons. As predicted, P1 pre-activation greatly enhanced the frequency of LC10-induced wing extensions (Figures 7F and 7G). These observations are consistent with the view that, in aroused males, visual information conveyed via LC10 cells is channeled toward the brain circuits that control courtship song.

DISCUSSION

Our work has identified the LC10 visual projection neurons as a major visual pathway mediating directed courtship in *Drosophila* males. LC10 neurons are broadly tuned to moving objects, which they detect via an antagonistic center-surround mechanism. If LC10 neurons are silenced, males are less able to track the female and select the ipsilateral wing to sing. Conversely, if LC10 neurons are unilaterally activated, isolated males turn ipsilaterally and extend the ipsilateral wing. The frequency of wing extensions is significantly enhanced if males

are placed into a state of arousal via pre-activation of the male-specific P1 neurons. LC10 neurons are morphologically and functionally similar in males and females, but the processing of their inputs in the brain is evidently both sex-specific and state-dependent. This processing likely also uses positional information encoded in the spatial pattern of LC10 activity to control the laterality of motor output.

The Role of LC10 Visual Projection Neurons in Directing Sex-Specific and State-Dependent Courtship Actions

Upon encountering another fly, a *Drosophila* male decides whether or not to court based on long range olfactory cues (Datta et al., 2008; Dweck et al., 2015; Kurtovic et al., 2007) and short range gustatory cues (Billeter et al., 2009; Lu et al., 2012; Thistle et al., 2012; Toda et al., 2012). These chemosensory cues are thought to signal the target's species, gender and receptivity. The role of vision in courtship initiation has been less clear. An earlier suggestion that flies might discriminate genders by visual assessment of body pigmentation (Kopp et al., 2000) would seem to be contradicted by the results of experiments that have manipulated pheromone profiles and clearly demonstrated that courtship initiation tracks with the pheromone profile, not the body pigmentation, of the target (Billeter et al., 2009; Demir and Dickson, 2005; Ferveur et al., 1997). Consistent with the view that visual cues have little bearing on the initial decision to court in our courtship chambers, we have shown here that blind males present wing extension frequencies similar to those of visually intact males. Nonetheless, an inanimate moving object the size of another fly can elicit courtship from a male (Agrawal et al., 2014), and unilateral activation of LC10 neurons in isolated males can induce ipsilateral wing extensions, albeit at a relatively low frequency. Thus, males might preferentially initiate courtship in response to chemical cues, as these more reliably signal gender and species identity, but in the absence of any chemosensory information can still be aroused, albeit less potently and specifically, by the visual detection of a fly-size moving object.

The chemosensory cues that trigger courtship are believed to activate the male-specific P1 neurons (Clowney et al., 2015; Kallman et al., 2015; Kohatsu et al., 2011), which induce a persistent state of courtship arousal (Poehlmann et al., 2014; Inagaki et al., 2014). In this state, visual tracking and directed wing extension are greatly enhanced. For example, either exposing the male foreleg to female contact pheromones (Kohatsu et al., 2011), or directly activating P1 neurons using optogenetics (Kohatsu and Yamamoto, 2015), stimulates a male to track a moving visual target. Similarly, an inanimate moving object is much more potent in eliciting courtship from a male if it is perfumed with female pheromones (Agrawal et al., 2014). Our data suggest that these conditional courtship actions are due to the sex-specific and state-dependent processing of visual information provided by LC10 neurons.

In larger flies, several classes of visual projection neurons have been proposed to function in the context of courtship pursuits performed by males during flight (Gilbert and Strausfeld, 1991; Nordström et al., 2006, 2008; Trischler et al., 2007). Typically, these small target feature neurons are sexually dimorphic and narrowly tuned to rather small, dark bars or spots and, in some cases, display directional selectivity. In contrast, *Drosophila* LC10 neurons are present in both sexes (Kimura et al., 2008; Wu et al., 2016) and are sensitive to bars with opposing contrasts and a broad range of angular sizes and speeds. Although they express *fru*, we have not noted any anatomical or functional differences between the sexes. The fru^{+} DA1 class of olfactory projection neurons that respond to the pheromone *cis*-vaccenyl acetate are similarly functionally monomorphic (Datta et al., 2008), but connect to distinct postsynaptic neurons in males and females (Kohl et al., 2013). This dimorphic connectivity is controlled by *fru* and likely contributes to the distinct behavioral responses of males and females to *cis*-vaccenyl acetate. The more compact termini of the LC10 visual projection neurons in the AOTu, together with the lack of any knowledge of their postsynaptic targets, makes it difficult at this point to discern whether *fru* instructs similar dimorphic connectivity downstream of LC10 neurons, but we speculate that this is likely to be the case.

While LC10 neurons do not appear to be the *Drosophila* analogs of the highly selective small target feature neurons of these larger flies, it remains possible that some other class of visual projection neurons fulfills this role in *Drosophila*. Perhaps the closest analog to these cells in *Drosophila* are the LC11 neurons, which, although not sexually dimorphic, are tuned to smaller objects than LC10 neurons and might therefore be more important in long range courtship initiation (Keleş and Frye, 2017). It is, however, also possible that the sexually dimorphic small target feature detection neurons of larger flies are a unique specialization associated with the challenge of relying primarily on visual cues for species and gender identification during flight—tasks for which *Drosophila* relies more on chemical than visual cues.

Extracting and Using the Visual Information Encoded by LC10 Neurons

A courting male detects the female as a moving object superimposed on a visual scene that itself may be moving on the retina due to the male's own locomotion. Our work suggests that LC10 neurons rely on a motion-based center surround to detect small objects against the visual background. In this mechanism, an LC10 cell is activated by the detection of movement within its receptive field but inhibited by movement just outside this field.

How might the information provided by LC10 neurons be used by circuits in the central brain to orient courtship behavior toward the female? In principle, three different types of algorithm might be employed to locate the female. First, the laterality of the target could be determined by comparing activities of left and right LC10 neurons. Second, because the LC10 termini are arranged topographically in the AOTu (Wu et al., 2016) and exhibit small receptive fields, they can also convey the precise angle toward the female. Third, the speed and direction of movement could be computed from the temporal pattern of LC10 activity across the topographic array of their termini in the AOTu, allowing downstream brain circuits to predict the female's future position. For ipsilateral wing extension, the first and simplest algorithm would suffice. An orienting turn, however, would likely involve the computation of a turning angle from the female's position and movement. Additionally, the

number of LC10 cells activated might also be used to estimate the distance to the female, which would determine the amplitude of any song produced, and whether or not the orienting turn is accompanied by an increase in forward locomotion.

In summary, our findings establish a link between visual feature detection and specific motor responses, identifying the LC10 visual projection neurons as moving object detectors used for directed courtship. This work paves the way for a detailed investigation of the neural circuits and computations that extract these visual features and use them in a sex-specific and state-dependent manner. Our functional characterization of LC10 visual projection neurons thus represents a critical first step in elucidating the neural processing that underlies a paradigmatic visuomotor transformation.

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Fly Genetics
  - Courtship Behavior
  - Automatic video tracking with MateBook
  - Behavior upon presentation of isolated visual input with the fly-on-ball assay
  - Unilateral activation using spatially restricted light
  - Artificial activation
  - Immunohistochemistry
  - 2-Photon calcium imaging
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of MateBook data for behavior assays in an open arena
  - Analysis of locomotion in assays with tethered males walking on a suspended ball
  - Analysis of 2-Photon calcium imaging
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and two tables and can be found with this article online at https://doi.org/10.1016/j.cell.2018.06.020.

ACKNOWLEDGMENTS

We are grateful to T. Liu, S. Prech, J. Haag, A. Leonhardt, and B. Arthur for expertise in various aspects of this work and to A. Arenz and A. Mauss for critically reading the manuscript. This work was supported by the Howard Hughes Medical Institute (HHMI) and Max Planck Gesellschaft (MPG).

AUTHOR CONTRIBUTIONS

I.M.A.R, B.J.D., and A. Borst conceived and designed this study. M.D. performed the functional imaging and data analysis presented in Figure 5. I.M.A.R. performed all other experiments and analyzed the data. A. Bahl and M.D. conceived and designed the visual stimuli. C.M. and B.J.D. created MateBook. I.M.A.R and B.J.D. co-wrote this paper with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 3, 2017
Revised: March 7, 2018
Accepted: June 10, 2018
Published: July 19, 2018

REFERENCES

Agrawal, S., Safarik, S., and Dickinson, M. (2014). The relative roles of vision and chemosensation in mate recognition of Drosophila melanogaster. J. Exp. Biol. *217*, 2796–2805.

Aptekar, J.W., Keleş, M.F., Lu, P.M., Zolotova, N.M., and Frye, M.A. (2015). Neurons forming optic glomeruli compute figure-ground discriminations in Drosophila. J. Neurosci. *35*, 7587–7599.

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.

Baden, T., Berens, P., Franke, K., Román Rosón, M., Bethge, M., and Euler, T. (2016). The functional diversity of retinal ganglion cells in the mouse. Nature *529*, 345–350.

Bahl, A., Ammer, G., Schilling, T., and Borst, A. (2013). Object tracking in motion-blind flies. Nat. Neurosci. *16*, 730–738.

Bahl, A., Serbe, E., Meier, M., Ammer, G., and Borst, A. (2015). Neural mechanisms for Drosophila contrast vision. Neuron *88*, 1240–1252.

Bath, D.E., Stowers, J.R., Hörmann, D., Poehlmann, A., Dickson, B.J., and Straw, A.D. (2014). FlyMAD: rapid thermogenetic control of neuronal activity in freely walking Drosophila. Nat. Methods *11*, 756–762.

Billeter, J.C., Rideout, E.J., Dornan, A.J., and Goodwin, S.F. (2006). Control of male sexual behavior in Drosophila by the sex determination pathway. Curr. Biol. *16*, R766–R776.

Billeter, J.C., Atallah, J., Krupp, J.J., Millar, J.G., and Levine, J.D. (2009). Specialized cells tag sexual and species identity in Drosophila melanogaster. Nature *461*, 987–991.

Boggs, R.T., Gregor, P., Idriss, S., Belote, J.M., and McKeown, M. (1987). Regulation of sexual differentiation in D. melanogaster via alternative splicing of RNA from the transformer gene. Cell *50*, 739–747.

Borst, A. (2014). Fly visual course control: behaviour, algorithms and circuits. Nat. Rev. Neurosci. *15*, 590–599.

Buchner, E. (1976). Elementary movement detectors in an insect visual system. Biol. Cybern. *24*, 85–101.

Budick, S.A., and O'Malley, D.M. (2000). Locomotor repertoire of the larval zebrafish: swimming, turning and prey capture. J. Exp. Biol. *203*, 2565–2579.

Bulthoff, H. (1982). Drosphila mutants disturbed in visual orientation. I. Mutants affected in early visual processing. Biol. Cybern. *45*, 63–70.

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.

Clowney, E.J., Iguchi, S., Bussell, J.J., Scheer, E., and Ruta, V. (2015). Multimodal chemosensory circuits controlling male courtship in Drosophila. Neuron *87*, 1036–1049.

Coen, P., Xie, M., Clemens, J., and Murthy, M. (2016). Sensorimotor transformations underlying variability in song intensity during Drosophila courtship. Neuron *89*, 629–644.

Collet, T. (1972). Visual neurones in the anterior optic tract of the privet hawk moth. J. Comp. Physiol. *78*, 396–433.

Cook, R. (1979). The courtship tracking of Drosophila melanogaster. Biol. Cybern. *34*, 91–106.

Cook, R. (1980). The extent of visual control in the courtship tracking of D. melanogaster. Biol. Cybern. *37*, 41–51.

Dankert, H., Wang, L., Hoopfer, E.D., Anderson, D.J., and Perona, P. (2009). Automated monitoring and analysis of social behavior in Drosophila. Nat. Methods *6*, 297–303.

Datta, S.R., Vasconcelos, M.L., Ruta, V., Luo, S., Wong, A., Demir, E., Flores, J., Balonze, K., Dickson, B.J., and Axel, R. (2008). The Drosophila pheromone cVA activates a sexually dimorphic neural circuit. Nature *452*, 473–477.

Del Bene, F., Wyart, C., Robles, E., Tran, A., Looger, L., Scott, E.K., Isacoff, E.Y., and Baier, H. (2010). Filtering of visual information in the tectum by an identified neural circuit. Science *330*, 669–673.

Demir, E., and Dickson, B.J. (2005). fruitless splicing specifies male courtship behavior in Drosophila. Cell *121*, 785–794.

Dhande, O.S., Stafford, B.K., Lim, J.A., and Huberman, A.D. (2015). Contributions of retinal ganglion cells to subcortical visual processing and behaviors. Annu. Rev. Vis. Sci. *1*, 291–328.

Dweck, H.K., Ebrahim, S.A., Thoma, M., Mohamed, A.A., Keesey, I.W., Trona, F., Lavista-Llanos, S., Svatoš, A., Sachse, S., Knaden, M., and Hansson, B.S. (2015). Pheromones mediating copulation and attraction in Drosophila. Proc. Natl. Acad. Sci. USA *112*, E2829–E2835.

Ferveur, J.F., Störtkuhl, K.F., Stocker, R.F., and Greenspan, R.J. (1995). Genetic feminization of brain structures and changed sexual orientation in male Drosophila. Science *267*, 902–905.

Ferveur, J.F., Savarit, F., O'Kane, C.J., Sureau, G., Greenspan, R.J., and Jallon, J.M. (1997). Genetic feminization of pheromones and its behavioral consequences in Drosophila males. Science *276*, 1555–1558.

Fujiwara, T., Cruz, T.L., Bohnslav, J.P., and Chiappe, M.E. (2017). A faithful internal representation of walking movements in the Drosophila visual system. Nat. Neurosci. *20*, 72–81.

Gahtan, E., Tanger, P., and Baier, H. (2005). Visual prey capture in larval zebrafish is controlled by identified reticulospinal neurons downstream of the tectum. J. Neurosci. *25*, 9294–9303.

Gilbert, C., and Strausfeld, N.J. (1991). The functional organization of male-specific visual neurons in flies. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *169*, 395–411.

Haikala, V., Joesch, M., Borst, A., and Mauss, A.S. (2013). Optogenetic control of fly optomotor responses. J. Neurosci. *33*, 13927–13934.

Hall, J.C. (1994). The mating of a fly. Science *264*, 1702–1714.

Inagaki, H.K., Jung, Y., Hoopfer, E.D., Wong, A.M., Mishra, N., Lin, J.Y., Tsien, R.Y., and Anderson, D.J. (2014). Optogenetic control of Drosophila using a red-shifted channelrhodopsin reveals experience-dependent influences on courtship. Nat. Methods *11*, 325–332.

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.

Kallman, B.R., Kim, H., and Scott, K. (2015). Excitation and inhibition onto central courtship neurons biases Drosophila mate choice. eLife *4*, e11188.

Keleş, M.F., and Frye, M.A. (2017). Object-detecting neurons in Drosophila. Curr. Biol. *27*, 680–687.

Kim, A.J., Fenk, L.M., Lyu, C., and Maimon, G. (2017). Quantitative predictions orchestrate visual signaling in Drosophila. Cell *168*, 280–294.

Kimura, K., Hachiya, T., Koganezawa, M., Tazawa, T., and Yamamoto, D. (2008). Fruitless and doublesex coordinate to generate male-specific neurons that can initiate courtship. Neuron *59*, 759–769.

Klapoetke, N.C., Murata, Y., Kim, S.S., Pulver, S.R., Birdsey-Benson, A., Cho, Y.K., Morimoto, T.K., Chuong, A.S., Carpenter, E.J., Tian, Z., et al. (2014). Independent optical excitation of distinct neural populations. Nat. Methods *11*, 338–346.

Kohatsu, S., and Yamamoto, D. (2015). Visually induced initiation of Drosophila innate courtship-like following pursuit is mediated by central excitatory state. Nat. Commun. *6*, 6457.

Kohatsu, S., Koganezawa, M., and Yamamoto, D. (2011). Female contact activates male-specific interneurons that trigger stereotypic courtship behavior in Drosophila. Neuron *69*, 498–508.

Kohl, J., Ostrovsky, A.D., Frechter, S., and Jefferis, G.S. (2013). A bidirectional circuit switch reroutes pheromone signals in male and female brains. Cell *155*, 1610–1623.

Kopp, A., Duncan, I., Godt, D., and Carroll, S.B. (2000). Genetic control and evolution of sexually dimorphic characters in Drosophila. Nature *408*, 553–559.

Krstic, D., Boll, W., and Noll, M. (2009). Sensory integration regulating male courtship behavior in Drosophila. PLoS ONE *4*, e4457.

Kurtovic, A., Widmer, A., and Dickson, B.J. (2007). A single class of olfactory neurons mediates behavioural responses to a Drosophila sex pheromone. Nature *446*, 542–546.

Land, M.F., and Collet, T.S. (1974). Chasing behaviour of houseflies (Fannia canicularis). J. Comp. Physiol. *89*, 331–357.

Lee, T., and Luo, L. (1999). Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. Neuron *22*, 451–461.

Lu, B., LaMora, A., Sun, Y., Welsh, M.J., and Ben-Shahar, Y. (2012). ppk23-Dependent chemosensory functions contribute to courtship behavior in Drosophila melanogaster. PLoS Genet. *8*, e1002587.

Luan, H., Peabody, N.C., Vinson, C.R., and White, B.H. (2006). Refined spatial manipulation of neuronal function by combinatorial restriction of transgene expression. Neuron *52*, 425–436.

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.

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.

Markow, T.A. (1987). Behavioral and sensory basis of courtship success in Drosophila melanogaster. Proc. Natl. Acad. Sci. USA *84*, 6200–6204.

Markow, T.A., and Hanson, S.J. (1981). Multivariate analysis of Drosophila courtship. Proc. Natl. Acad. Sci. USA *78*, 430–434.

Markow, T.A., and Manning, M. (1980). Mating success of photoreceptor mutants of Drosophila melanogaster. Behav. Neural Biol. *29*, 276–280.

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.

Mellert, D.J., Knapp, J.M., Manoli, D.S., Meissner, G.W., and Baker, B.S. (2010). Midline crossing by gustatory receptor neuron axons is regulated by fruitless, doublesex and the Roundabout receptors. Development *137*, 323–332.

Nern, A., Pfeiffer, B.D., and Rubin, G.M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. Proc. Natl. Acad. Sci. USA *112*, E2967–E2976.

Nikonov, S.S., Kholodenko, R., Lem, J., and Pugh, E.N., Jr. (2006). Physiological features of the S- and M-cone photoreceptors of wild-type mice from single-cell recordings. J. Gen. Physiol. *127*, 359–374.

Nordström, K., Barnett, P.D., and O'Carroll, D.C. (2006). Insect detection of small targets moving in visual clutter. PLoS Biol. *4*, e54.

Nordström, K., Barnett, P.D., Moyer de Miguel, I.M., Brinkworth, R.S., and O'Carroll, D.C. (2008). Sexual dimorphism in the hoverfly motion vision pathway. Curr. Biol. *18*, 661–667.

Olberg, R.M., Worthington, A.H., and Venator, K.R. (2000). Prey pursuit and interception in dragonflies. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *186*, 155–162.

Otsuna, H., and Ito, K. (2006). Systematic analysis of the visual projection neurons of Drosophila melanogaster. I. Lobula-specific pathways. J. Comp. Neurol. *497*, 928–958.

Pan, Y., Meissner, G.W., and Baker, B.S. (2012). Joint control of Drosophila male courtship behavior by motion cues and activation of male-specific P1 neurons. Proc. Natl. Acad. Sci. USA *109*, 10065–10070.

Pfeiffer, B.D., Ngo, T.T., Hibbard, K.L., Murphy, C., Jenett, A., Truman, J.W., and Rubin, G.M. (2010). Refinement of tools for targeted gene expression in Drosophila. Genetics *186*, 735–755.

Pologruto, T.A., Sabatini, B.L., and Svoboda, K. (2003). ScanImage: flexible software for operating laser scanning microscopes. Biomed. Eng. Online *2*, 13.

Portugues, R., Feierstein, C.E., Engert, F., and Orger, M.B. (2014). Whole-brain activity maps reveal stereotyped, distributed networks for visuomotor behavior. Neuron *81*, 1328–1343.

Reichardt, W., and Poggio, T. (1976). Visual control of orientation behaviour in the fly. Part I. A quantitative analysis. Q. Rev. Biophys. *9*, 311–375, 428–338.

Sakai, T., Isono, K., Tomaru, M., and Oguma, Y. (1997). Light-affected male following behavior is involved in light-dependent mating in Drosophila melanogaster. Genes Genet. Syst. *72*, 275–281.

Scavarda, N.J., O'tousa, J., and Pak, W.L. (1983). Drosophila locus with gene-dosage effects on rhodopsin. Proc. Natl. Acad. Sci. USA *80*, 4441–4445.

Schnell, B., Raghu, S.V., Nern, A., and Borst, A. (2012). Columnar cells necessary for motion responses of wide-field visual interneurons in Drosophila. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *198*, 389–395.

Seelig, J.D., Chiappe, M.E., Lott, G.K., Dutta, A., Osborne, J.E., Reiser, M.B., and Jayaraman, V. (2010). Two-photon calcium imaging from head-fixed Drosophila during optomotor walking behavior. Nat. Methods *7*, 535–540.

Sweeney, S.T., Broadie, 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.

Thistle, R., Cameron, P., Ghorayshi, A., Dennison, L., and Scott, K. (2012). Contact chemoreceptors mediate male-male repulsion and male-female attraction during Drosophila courtship. Cell *149*, 1140–1151.

Tirian, L., and Dickson, B.J. (2017). The VT GAL4, LexA, and split-GAL4 driver line collections for targeted expression in the Drosophila nervous system. bioRxiv. https://doi.org/10.1101/198648.

Toda, H., Zhao, X., and Dickson, B.J. (2012). The Drosophila female aphrodisiac pheromone activates ppk23(+) sensory neurons to elicit male courtship behavior. Cell Rep. *1*, 599–607.

Trischler, C., Boeddeker, N., and Egelhaaf, M. (2007). Characterisation of a blowfly male-specific neuron using behaviourally generated visual stimuli. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *193*, 559–572.

von Philipsborn, A.C., Liu, T., Yu, J.Y., Masser, C., Bidaye, S.S., and Dickson, B.J. (2011). Neuronal control of Drosophila courtship song. Neuron *69*, 509–522.

von Reyn, C.R., Nern, A., Williamson, W.R., Breads, P., Wu, M., Namiki, S., and Card, G.M. (2017). Feature integration drives probabilistic behavior in the Drosophila escape response. Neuron *94*, 1190–1204.

Wu, M., Nern, A., Williamson, W.R., Morimoto, M.M., Reiser, M.B., Card, G.M., and Rubin, G.M. (2016). Visual projection neurons in the *Drosophila* lobula link feature detection to distinct behavioral programs. eLife *5*, e21022.

Yu, J.Y., Kanai, M., Demir, E., Jefferis, G.S., and Dickson, B.J. (2010). Cellular organization of the neural circuit that drives Drosophila courtship behavior. Curr Biol. *20*, 1602–1614.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-Bruchpilot antibody (nc82, monoclonal, mouse)	Developmental Studies Hybridoma Bank	RRID: AB_2314866
anti-GFP antibody (chicken)	Rockland	Cat#600-901-215S; RRID: AB_1537403
anti-RFP antibody (rabbit)	Rockland	Cat#600-401-379; RRID: AB_11182807
anti-FLAG antibody (rat)	Novus Biologicals	Cat# NBP1-06712; RRID: AB_1625981
Goat anti-chicken, Alexa488	Thermo Fisher Scientific	Cat#A-11039; RRID: AB_2534096
Goat anti-mouse, Alexa633	Thermo Fisher Scientific	Cat#A-21050; RRID: AB_2535718
Goat anti-rabbit, Alex568	Thermo Fisher Scientific	Cat#A-11011; RRID: AB_143157
Chemicals, Peptides, and Recombinant Proteins		
Paraformaldehyde (PFA)	Electron Microscopy Sciences	Cat# 15713
Schneider's medium	Sigma	Cat#S0146
Normal goat serum	Sigma	Cat#G9023
Experimental Models: Organisms/Strains		
Drosophila: Canton S	N/A	N/A
Drosophila: ort^{1}	Bloomington Drosophila Stock Center	Cat#BL1133
Drosophila: ninaE^{7}	Bloomington Drosophila Stock Center	Cat#BL1613
Drosophila: +; UAS-TNT; +	Sweeney et al., 1995	N/A
Drosophila: +; UAS-IMPTNT-Q: +	Sweeney et al., 1995	N/A
Drosophila: w^{1118}*; R59E08.p65ADZp; R42F06.ZpGAL4DBD*	Bahl et al., 2015	N/A
Drosophila: w^{1118}*; VT040012.p65ADZp; VT043656.ZpGAL4DBD*	this study	N/A
Drosophila: w^{1118}*; +; VT043656-GAL4*	this study	N/A
Drosophila: w^{1118}*; R35D04.p65ADZp; VT043656.ZpGAL4DBD*	this study	N/A
Drosophila: w^{1118}*; R25H09.p65ADZp; VT043656.ZpGAL4DBD*	this study	N/A
Drosophila: w; UAS-GCaMP6m @ attP40	Chen et al., 2013	N/A
Drosophila: w; UAS-CD8:Tomato; +	Toda et al., 2012	N/A
Drosophila: w; UAS-MCFO	Nern et al., 2015	N/A
Drosophila: w, UAS-FRT-STOP-FRT-CsChrimson:Venus, hs-FLP2:PEST;+;+	Wu et al., 2016	N/A
Drosophila: +, UAS-CsChrimson:Venus;+;+	Wu et al., 2016	N/A
Drosophila: w; VT029314-lexA;+	this paper	N/A
Drosophila: w, LexAop-CsChrimson:Venus;+;+	Gerald Rubin	N/A
Drosophila: +; NP2361-GAL4, UAS-FRT-STOP-FRT-TrpA1:myc; fru^{FLP}	von Philipsborn et al., 2011	N/A
Drosophila: w; NP2361-GAL4, UAS-FRT-STOP-FRT-CD8:GFP; fru^{FLP}	von Philipsborn et al., 2011	N/A
Drosophila: w; VT043920.p65ADZp; R91B01.ZpGAL4DBD	Wu et al., 2016	N/A
Drosophila: w; R35D04.p65ADZp; R22D06.ZpGAL4DBD	Wu et al., 2016	N/A
Drosophila: w; R72C08.p65ADZp; R50D07.ZpGAL4DBD	Wu et al., 2016	N/A
Drosophila: w; R35D04.p65ADZp; R80E07.ZpGAL4DBD	Wu et al., 2016	N/A
Drosophila: w; R72C08.p65ADZp; VT040747.ZpGAL4DBD	Wu et al., 2016	N/A
Drosophila: w; R35D04.p65ADZp; R91B01.ZpGAL4DBD	Wu et al., 2016	N/A

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Drosophila: w; VT040747.p65ADZp; R72C08.ZpGAL4DBD	Wu et al., 2016	N/A
Drosophila: w; R22H02.p65ADZp; R20G06.ZpGAL4DBD	Wu et al., 2016	N/A
Drosophila: +; UAS-tra; +	Ferveur et al., 1995	N/A
Software and Algorithms		
MateBook	this study	https://github.com/Dicksonlab/MateBook
custom-written software using Python 2.7	this study	https://github.com/Dicksonlab/MateBook
ScanImage	Pologruto et al., 2003	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Barry J. Dickson (dicksonb@janelia.hhmi.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Several lines of the fruit fly *Drosophila melanogaster* were used in this study and are listed in Table S2 and Key Resources Table. Flies were raised in vials containing standard cornmeal-agar medium supplemented with baker's yeast and incubated at 25°C with 60% humidity and 12h bright light/dark cycle throughout development and adulthood. Males were the subject of interest in all experiments, with the exception of functional imaging in Figures S6D–S6O. In this case, females were treated similarly to males (see below). For single-pair courtship behavior assays (Figures 1, 2, 6, S1, S2, S3, S5, and S6), males were collected as virgins (i.e., 2 to 8 hours post-eclosion), reared in isolation and aged until 7 to 9 days old. Wild-type females used in the assay were collected as virgins, group housed and aged to 2 to 5 days old. Courtship assays were carried out within 3 hours after onset of light or before onset of darkness, respecting fly circadian peaks of activity. For experiments on the ball set up (Figures 3 and S5), males were group housed with other males and females of the same genotype and aged to 6 to 7 days old. Males used in functional imaging (Figures 4, 5, 6, and S6) were either group housed with females of the same genotype or reared in isolation and aged to 9 to 11 days old. Since no appreciable differences between group housed and isolated males were observed, the data from both conditions were pooled and analyzed together. For artificial activation experiments (Figures 7 and S7) males were group housed with other males of the same genotype and reared until 6 to 12 days old. For stochastic activation experiments (Figures 7 and S7), flies were heat-shocked at 1st-instar larva stage at 37°C for 15 min to 30 min, then returned to 25°C. For stochastic and full expression activation experiments, artificial activation relied on expression of CsChrimson, a rhodopsin-like cation channel activated by red light (Klapoetke et al., 2014). *all-trans retinal* (30mg/ml) was added to approx. 40ul yeast paste which replaced baker's yeast in otherwise unaltered fly food. Males were transferred to fresh *all-trans retinal* food vials every two days during adulthood and aged until 6 to 8 days old for stochastic expression and 8 to 12 days for full expression. Starting from adulthood, males were also reared under blue light until test day, with other incubation conditions unaltered. Males were incubated for 2 min at 31°C in order to achieve acute activation of P1 neurons in Figures 7F and 7G with *NP2361-GAL4*; light stimulation of LC10 neurons was performed after this incubation period, at 25°C.

METHOD DETAILS

Fly Genetics

The split lines *LC10-SS1*, *LC10-SS2* and *LC10-SS3* lines expressing in LC10 neurons were based on an initial behavioral screen in which *VT-GAL4* lines expressing in *fruitless*-positive visual projection neurons (T. Liu and B.J.D., unpublished observations) (Tirian and Dickson, 2017) were tested in single-pair courtship assays. The respective split-halves were generated by cloning the corresponding tiles into vectors containing either GAL4 split half (Pfeiffer et al., 2010) and further tested for expression in LC10 neurons. Other flies used in this study include: *Canton S*; *ninaE^{7}* mutants (Scavarda et al., 1983; BL2103); *ort^{1}* mutants (Bulthoff, 1982; BL1133); *UAS-TNT* (Sweeney et al., 1995); *UAS-CD8:GFP* (Lee and Luo, 1999); *UAS-CD8-tdTomato* (Toda et al., 2012); *UAS-GCaMP6m* in attP40 (Chen et al., 2013); *UAS-CsChrimson-Venus* in attP18 (Klapoetke et al., 2014); *UAS-MCFO* (Nern et al., 2015); *UAS-FRT-stop-FRT-CsChrimson-Venus* in attP18, *hsFlp2:PEST* in attP3 and the LC10 lines *OL19B*, *OL20B*, *OL22B*, *OL23B*, *SS3822*, *SS2669* and *SS2681* (Wu et al., 2016); *UAS-tra* (Ferveur et al., 1995); *NP2631-GAL4*, *UAS > stop > TrpA1:myc* and *UAS > stop > CD8:GFP* (von Philipsborn et al., 2011); *fru^{FLP}* (Yu et al., 2010); *w, LexAop-CsChrimson:Venus* at attP18 was a generous gift from Gerald Rubin. The complete genotypes used in each figure are listed in Table S2.

Courtship Behavior

To test courtship behavior, a naive male (reared in isolation) was paired with a virgin female inside a chamber measuring 18mm in diameter and 4mm in height, with a removable strip separating the chamber in two halves, in which each fly awaited the start of the assay. A single plate containing eight such chambers was placed roughly 20 cm under JVC HD video camera and illuminated with abundant bright light. The assay was initiated by removing the strip between male and female and was video-recorded with a resolution of 1440x1080 (16:9) at 25 fps for 10 min. Videos were acquired in the .MTS file format, without rendering, tracked with a custom software called MateBook (see below) and behavior classifiers as well as locomotion parameters were used for data analysis.

Automatic video tracking with MateBook

To quantify male and female movements as well as courtship steps objectively, we employed a video-tracking software called MateBook, which fits an ellipse to the segmented image of each fly and provides XY coordinates of the ellipse center and head position throughout the video. Gender identity was correctly attributed based on size difference (see Figure S1), since females are typically slightly larger than males. Its accuracy in resolving occlusions depends on the assay and video format. For all the assays reported here, we used chambers of 18mm diameter and acquired video recordings at 25 fps. The courtship steps were identified with the help of classifiers based on Dankert et al. (2009):

- Circling: the male keeps a minimum distance of 2 mm while walking sideways with a minimum speed of 3mm/s around the female; persistence is 0.5 s.
- Copulating: occlusion persisting for at least 45 s.
- Following: the male keeps distance to the female between 2 and 5mm, with maximum change of 2mm/s, while directly behind the female (absolute angle theta not larger than 30°), while both flies are walking with a minimum speed of 2mm/s; persistence is 1.0 s.
- Orienting: the male keeps the female in front (maximum absolute angle theta not larger than 60°) while walking at maximum 1mm/s; persistence is 1.0 s.
- Wing extension: the angle formed by the major axis of the ellipse fitted to the body and the line through the center of the ellipse and the tip of the wing exceeds 30°; persistence is 0.5 s.

Behavior upon presentation of isolated visual input with the fly-on-ball assay

To test male locomotion responses to specific visual stimuli isolated from chemosensory input, we used the fly-on-ball assay (Bahl et al., 2013; Buchner, 1976; Seelig et al., 2010). A spot of UV glue was applied to the top of the head and anterior thorax of a cold anesthetized male and a metal tether was carefully placed on the glue spot such that it united the head to the thorax. Blue light was applied for 5 s to curate the glue. Next, the wings of the male were glued to the tether to prevent flight from occurring during the assay. After 10 min recovery, the tethered male was then mounted on top of a 6mm foam sphere smoothly suspended by air flowing from a sphere holder. The air was slowly heated to 34°C during the first 10 min of each experiment (Bahl et al., 2013). Two optical tracking sensors were used to track the movement of the sphere from which fly locomotion was computed (Bahl et al., 2013). The visual stimuli were on 3 LCD screens placed around the fly, forming a visual arena that covered 135° azimuth and 57° elevation of the fly visual field and were designed with Pandas 3D and Python 2.7. The visual stimuli where:

- Open loop grating: whole field grating with a spatial wavelength of 20°, 60% Weber contrast moving at 20°/s for 2 s per 7 s trial was presented in open loop (Bahl et al., 2013).
- Open loop long bar: a long, vertical bar measuring 10° wide and spanning the elevation of the visual arena was moved at 60°/s, counter-clockwise or clockwise, starting from −60° (counter-clockwise) or 60° (clockwise) azimuthal position at Weber contrast of 100% for dark bar and 70% for bright bar. Each trial lasted 7 s.

The open loop grating and the open loop long bar, each in two directions, were presented in the same session to each fly. Together, each fly was exposed to 40 combinations of open loop stimuli with a stable temperature of 34°C – and thus constant walking - in the last 20 trials, which were used for analysis.

- Virtual cylinder with 3D bead: a virtual cylinder with 10cm height and 5cm radius was generated and lined with horizontal stripes bearing 20% Weber contrast. A virtual bead was placed inside this cylinder by creating an additional elliptic cylinder with half axes 0.25cm and 0.35cm and 0.25cm height. The virtual bead was made to describe a circle with a radius of 2.5cm at 0.05Hz (i.e., forward speed of 7.85mm/s) and was presented moving either in clockwise or counter-clockwise direction. At the start of each 60 s trial the male was randomly placed at any position in this virtual cylinder and the image presented on the visual arena was that of a camera located at the same XY coordinated as the center of the male. The cylinder walls were infinitively far away and the virtual bead was made to disappear from the screen if the male is close enough to touch the bead. The virtual bead was presented at two opposing contrasts with Weber contrast of 125%. In a third condition, the bead was made invisible such that distance and angle male to bead were still measured, but the male never visualized the bead. Each trial lasted 60 s.

Combinations of dark and bright beads, each in two directions, together with invisible bead lasted 180 s and were presented to the same fly for a total of fifteen times of which the last 12 (temperature stably at 34°C) were used for data analysis.

- Closed loop long bar: a single black 10° wide and 114° long bar was presented in closed loop exactly as in Bahl et al., 2013.

Unilateral activation using spatially restricted light

A tethered male placed on an air-suspended ball was exposed to a dark visual arena and to a 0.9mm diameter laser spot from a 635nm laser (LDM635/5LJM from Roithner LaserTechnik) placed roughly 30cm above the tethered male. The laser spot was touching the tip of either the left or right eye such that when light intensity was increased during stimulation, only half of the head was illuminated. 'Laser on' trials consisted of 10 s with 635nm laser coming on for 0.5 s with a light intensity of 32 $\mu W/mm^2$ and alternated with same duration 'laser off' trials. Twenty such trials per fly were used for data analysis.

Artificial activation

To obtain unilateral expression of CsChrimson in LC10 neurons we used stochastic activation based on a heat shock promotor driving the flipase Flp2:PEST stochastically to remove a stop cassette that prevented expression of CsChrimson (Wu et al., 2016). Flies were heat shocked at 37°C for 15 or 30 min at the first larva stage, otherwise raised at 25°C. Single adult males *LC10-SS1 > FRT-dSTOP-FRT-CsChrimson:Venus, hs-Flp2:PEST*, aged 6 to 8 days old, were placed in the same chambers used for courtship assays. The chambers were illuminated with blue light from below and video recorded with JVC HD cameras equipped with a red filter to reduce overexposure to red light during stimulation in the video. Activation trials consisted of red LED light of intensity 3.8mW/cm^2 for 0.5, 2, or 10 s, repeated five times. An extra blue LED light located to the side of the chamber turned on with red LEDs and allowed matching of the behavioral data with stimulation trials. Video recording and data analysis was similar to above with the exception that wing behavior was scored manually in which the scorer was blind to the genotype. The expression of CsChrimson in LC10 neurons was determined after the activation assay with dissection of the brain of each male. The expression pattern (no expression, unilateral and bilateral) was used to pool data from same expression flies for further analysis.

Activation of all LC10 neurons with *LC10-SS2* or *LC10-lexA* was performed in the same set up with the specified genotypes (see Table S2). Activation of P1 neurons was achieved by incubating males *NP2361-GAL4>UAS-TrpA1* at 31°C for 2 min.

Immunohistochemistry

Staining of the fly brain and ventral nerve chord was performed as previously described (Yu et al., 2010). Briefly brains and ventral nerve chords from 5 to 7 days old group-housed males were dissected in phosphate buffered saline (PBS; Figures 2A, 2B, and S3B) or in Schneider's insect media (all other figures) and fixed in 4% paraformaldehyde for 25 to 45min at room temperature (approx. 21 to 22°C). The brains were washed three times in 5 to 10min washes in PBST (PBS with 0.5% Triton X-100) followed by blocking in 10% normal goat serum in PBST for 2hrs at room temperature or overnight at 4°C. The primary and secondary antibodies were diluted in 5% normal goat serum in PBST and were incubated for 48 to 72h at 4°C. Several 15min washes in PBST at room temperature were performed between primary and secondary antibodies and before mounting the slide. Vectashield was used as the mounting media. Brains and ventral nerve chords were imaged using a Zeiss LSM 510 with a 25x oil immersion objective (Zeiss Multi Immersion Plan NeoFluar 25x/0.8) or a Leica SP5 with a 20x oil immersion objective (Leica Immersion 20x/0.7). See Key Resources Table for a list of antibodies used.

2-Photon calcium imaging

Calcium responses of LC10 neurons were recorded in w^- (Figures 4, 5, 6, and S6) and w^+ (Figure S6) backgrounds. The eye color of *LC10-SS1>GCaMP6m* males was similar in w^- and w^+ backgrounds (Figure S6P) as was the motion-based center surround (Figures 5 and S6). Flies were prepared as in Arenz et al. (2017). Briefly, a cold anesthetized fly was attached to a plastic column by the thorax with UV glue. The legs were spread out and glued to the same column with bee's wax. The head was bent ventrally such that the proboscis touched the thorax ventrally between the forelegs and midlegs and glued in this position with bee's wax, with care to leave the eyes and area in front of eyes free of wax. The column was mounted on an aluminum foil container, in turn built on a holder mountable on the top of the visual arena. The aluminum container was filled with Ringer solution (103mM NaCl, 3mM KCl, 5mM TES, 10mM trehalose, 10mM glucose, 3-7mM sucrose, 26mM $NaHCO_3$, 1mM NaH_2PO_4, 1.5 $CaCl_2$ and 4mM $MgCl_2$, pH 7.3 to 7.35, 280-290mOsmol/Kg) (Mauss et al., 2014). A hole was cut in the dorsal cuticle behind the eye, trachea and fat body were removed. The fly preparation was checked for presence of pumping in the trachea (over roughly 80% of preparations) and mounted on the visual arena.

We used a custom-built 2-photon microscope (Maisak et al., 2013), and the ScanImage software (Pologruto et al., 2003) for imaging and acquisition of calcium signals. Images were acquired at a resolution of 128x128 or 128x64 pixels and at a frame rate of 5.96 or 11.89 Hz, depending on the stimulus, as described below. A second channel was used to receive a trigger signal from the stimulus arena that marked the starting time point of each stimulus and hence allowed for synchronization of the signals with the stimuli shown.

Visual stimuli were presented on a custom-built projector based arena (Arenz et al., 2017). Briefly, two commercial micro-projectors were used to render the stimuli on an opaque, cylindrical rear projection screen that covered 180° in azimuth and approximately

105° in elevation of the visual field of the fly. The fly was placed in the center of that cylinder. Stimuli were shown using only the green LED (OSRAM L CG H9RN) covering wavelengths between 500nm and 600nm and with an image refresh rate of 180Hz. The maximum luminance of the arena is 276 ± 48 cd/m^2 if the whole screen is set to white (i.e., a grayscale value of 255). Stimuli were programmed using Python 2.7 and Panda3D, an open source 3D rendering framework for Python and C++ users. Custom written software predistorted the images such that the projected images appeared regularly on the curved arena screen.

Stimuli: For gratings, moving long or short bars, counter phase flickers, full-field flickers and edges, a texture covering the respective spatial parameters of the stimulus was placed on a virtual cylinder whose image was then projected onto the real cylindrical arena screen. Bright features in the stimuli had an intensity of 150 out of 255 grayscale values, and for dark features an intensity value of 25 was used. For looming spheres, the stimulus consisted of the image of a virtual sphere actually approaching the camera with a given constant velocity. This is equivalent to looming stimuli in which the size of the outline of a sphere is increased proportionally to the inverse tangent of the ratio between radius and distance of the sphere. The sphere had an initial size of 11.4° and approached within either 5 or 25 s. The image acquisition frame rate was 5.96 Hz.

For the motionless white noise, a randomly generated binary image of black and white squares covering 5° in azimuth and correspondingly in elevation was projected onto the arena screen and changed every 3 s. During each stimulus presentation 300 random images were shown.

For the motion noise stimulus, a pre-rendered video file (exported in H.264 format with lossless compression and 60Hz frame rate) was generated for later presentation on the arena screen. The screen was parted in 10 stripes and in each stripe a sine wave grating with 20° spatial wavelength, random initial phase and intensity scaling between 0 and 100 on the 8-bit grayscale value of the display was shown (Figure 5C). The expected result from a stimulus-response reverse correlation with such a stochastic motion stimulus would be zero for motion-selective neurons without direction selectivity, as is the case for LC10 neurons. We therefore restricted the velocities of the gratings to be front-to-back (progressive) in order to avoid averaging out correlations with the stimulus velocity due to this nonlinearity of the LC10 response. After stimulus onset, these 10 sine wave gratings moved all to one direction (front-to-back) but following a randomly generated velocity profile independently from each other. This velocity profile was generated before by generating normally distributed random numbers for each frame of the video, setting the variance of the distribution to 2 cycles/s and the mean value to 0.5 cycles/s. The resulting array was low-pass filtered with a Gaussian window so as to smooth out frequencies above 5 Hz. This low-pass filtered random trace was taken as a velocity profile to control the phase velocity of the sine wave grating at each time point (Figures 5D–5F). The mean of the phase velocity distribution, 0.5 Hz, corresponds for a 20° wavelength grating to an absolute velocity of 10°/s. Velocity values below zero were not allowed. The whole stimulus was 10 minutes long. For motion noise the image acquisition frame rate was 11.89 Hz whereas for other stimuli we used 5.96 Hz.

To test whether the temporal properties of the motion noise stimulus (phase velocities of the sine waves restricted below approximately 2 cycles/s) had a severe impact on our estimate of the neuron's temporal filtering properties, we performed a Wiener deconvolution on the data. To this end, the temporal component of the stimulus-response cross-correlation function was taken as the cross-section through the peak of the spatiotemporal cross-correlation function. We then calculated the Fourier-transform of the temporal component and estimated a frequency-dependent noise function using the standard deviation of that function over flies. We then applied a standard Wiener deconvolution algorithm to remove the impact of the non-white power-spectrum of our stimulus on our measurement (Figures S6U–S6Z).

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of MateBook data for behavior assays in an open arena

MateBook was set to track two flies for single-pair courtship assay and provides parameters on a per frame and per fly-pair basis, which were uploaded into Python2.7 for further processing. To plot distance, angle θ and relative position distributions, frames identified as misssegmented (mseg), occlusion (occ) or copulation (cop) by MateBook were removed from further analysis. For distance and angle θ distributions, a probability density function was calculated per fly and averaged over flies. For female relative position heatmaps, all frames identified by MateBook as male extending a single wing (we [0]) were compiled. MateBook also provided classification of behavior on a per frame and per fly-pair basis using classifiers partially based on Dankert et al., 2009 but further refined (see above). Behavioral raster plots (Figures S1 and S3) represented as ethograms relied on the per frame, per fly-pair classifier information with the hierarchy: orienting > following > wing extension > circling > copulation.

With the exception of wing choice and copulation, indices were calculated per fly-pair as:

$$\text{behavior index} = \text{number of frames with behavior} / \text{total number of frames till copulation of end of movie} * 100.$$

To compare following, orienting or wing extension indices across genotypes, P values were calculated using the Student's t test with Bonferroni correction.

The copulation index was calculated as:

$$\text{Copulation index} = \text{number of pairs copulating at the end of 10 min} / \text{total number of pairs tested} * 100.$$

To compare copulation indices across genotypes, P values were calculated with Fisher's exact test with Bonferroni correction.

A frame with single wing extension (we) was considered ipsilateral if the angle θ male to female was between −15° and 90° if the extended wing was the right wing and between −90° and 15° if the extended wing was the left wing (where a positive angle θ means right side). The wing extension was considered contralateral in if angle θ was outside this interval. To quantify the wing bias for single wing extensions (WE in formula), a wing choice index was computed per fly:

$$\text{wing choice index} = \frac{\text{number of frames ipsilateral WE} - \text{number of frames contralateral WE}}{\text{number of frames ipsilateral WE} + \text{number of frames contralateral WE}}.$$

The angle the female subtended on the male retina, i.e., angular size in the azimuth, was computed as the projection of an ellipsoid body, with major and minor axis based on the ellipse fitted on the female, onto the point with the XY coordinates of the male head per frame. In other words, the angular size of the female was the angle of the triangle vertex that sits on the male head XY coordinates. The triangle was drawn from the lines connecting the male head XY coordinates with the two points at the limits of the ellipse. The location of these two poins on the ellipse was determined by the projection of the ellipsoid body onto the XY coordinates of the male head.

For analysis of open arena artificial activation assays, MateBook was set to track one fly per arena. Frames labeled as misssegmented (mseg) were infrequent and were not removed. Frames with stimulus light on were acquired from a blue LED located laterally to the chambers which turned on with red light. Speed traces were plotted starting 10 s before the onset of red stimulus light (spaced every 30 or 60 s) and speed data per trial were pooled and averaged per fly and then averaged over all flies with the same genotype and expression. Because the base light that backlit the arena in activation assays was dim and because isolated males tended to spend time grooming their wings, wing behavior had to be manually scored (data in Figures 7E–7G). This scoring was blind to expression pattern (Figure 7E) or genotype (Figures 7F and 7G). The frequency of wing extensions were compared across genotypes by calculating P values using the Student's t test with Bonferroni correction.

Analysis of locomotion in assays with tethered males walking on a suspended ball

Over all behavioral tests on the ball set up, males were discarded if the average forward speed over all trials at 34°C was below 3mm/s or the average absolute rotational speed was higher than 30°/s. Forward and rotational speeds, as well as angle and distance male to bead when applicable (virtual cylinder with 3D bead and closed loop long bar), were first averaged over same-type trials and then over flies preserving trial structure. For tests with the virtual cylinder and 3D bead, for each frame the position of the 3D virtual bead relative to the male was computed from the distance and angle male to the bead using the cosine for ΔY and sine for ΔX. Frames of same type trials were concatenated over all flies and used to plot the heatmaps in Figures 3H–3S and S5J–S5T.

Analysis of 2-Photon calcium imaging

All data analysis was performed using custom-written software using Python 2.7 and taking advantage of the Python libraries SciPy 0.16.1, OpenCV 2.4.8 and NumPy 1.10.2. Artifacts originating from the movement of the sample while imaging were corrected for by automatically registering the calcium imaging stacks across time. With the exception of calcium responses to noise stimuli, changes in GCaMP6m fluorescence were analyzed as follows: regions of interest were selected based on signal correlation over time of each pixel with its six neighboring pixels and a threshold (Portugues et al., 2014) or were manually drawn on single punctae present in LC10 dendrites in the lobula or terminals in the anterior optic tubercle (AOTu).

Over all calcium imaging sessions, a single bright square-bar 20° was presented moving at 70°/s at several elevations to determine the preferred elevation of the current preparation. Several stimuli were combined into protocols in a pseudo-random order and presented three times per acquisition. The baseline calcium response F0 was the average of GCaMP fluorescence measured in the 15 frames prior to onset of the first stimulus in the acquisition. ΔF/F0 was computed and normalized per ROI over each acquisition. Triggers were then used to group the repetitions per stimulus and average them per fly and subsequently over flies.

For motionless white noise and motion noise, first a ''dynamic mean signal'' was calculated by low-pass filtering the signal during stimulus presentation with a Gaussian window of 120 s standard deviation. The resulting trace was subtracted from the signal to remove the offset and slow baseline fluctuations from the signal. Then the stimulus-response reverse correlation function was calculated. For motionless white noise, the signal was parted in 300 bins, according to the time frame (3 s long) of the presentation of each of the 300 presented random images. The mean response R_n was calculated for each of these bins frames. Then, the stimulus-response correlation function is given by

$$K(x,\ y) = \sum_n S_n(x,y) \cdot R_n,$$

where $S_n(x,y)$ denotes the stimulus (depending on both spatial dimensions). The resulting weighted average $K(x,y)$ corresponds to the spatial correlation of the signal with this static white noise stimulus, i.e., the spatial receptive field (as estimated by this stimulus). When reverse correlation yielded defined and visible spatial receptive fields, the receptive field centers were determined by fitting a Gaussian to each of the cross-sections through the peak along both space axes. Then, all spatial receptive fields obtained were shifted upon each other and after normalization to the peak a mean spatial receptive field across all acquisitions was calculated (see Figure 5D).

For motion noise, the reverse correlation was continuous in time. This stimulus is 1-dimensional and given by the random velocity profile $v(y,t)$ (depending on the vertical coordinate y and time t) that controlled the motion of each of the 10 sine wave gratings. Here, the stimulus-response reverse correlation function is

$$K(y,\tau) = \int_0^T dt\, v(y,t-\tau) \cdot R(t).$$

The result is a spatio-temporal receptive field, that indicates the correlation of the signal with the velocity of each of the 10 gratings. The spatio-temporal receptive fields were normalized in z-score and only receptive fields with peak amplitudes above 7 standard deviations from the mean were taken (to avoid noise artifacts). Spatio-temporal receptive fields from different acquisitions and ROI's were peak-aligned and a mean spatio-temporal receptive field was calculated (Figures 5G–5I).

DATA AND SOFTWARE AVAILABILITY

MateBook software together with instructions for the user and the developer are available through GitHub, together with the Custom-written software using Python 2.7 used to analyze MateBook data: https://github.com/Dicksonlab/MateBook

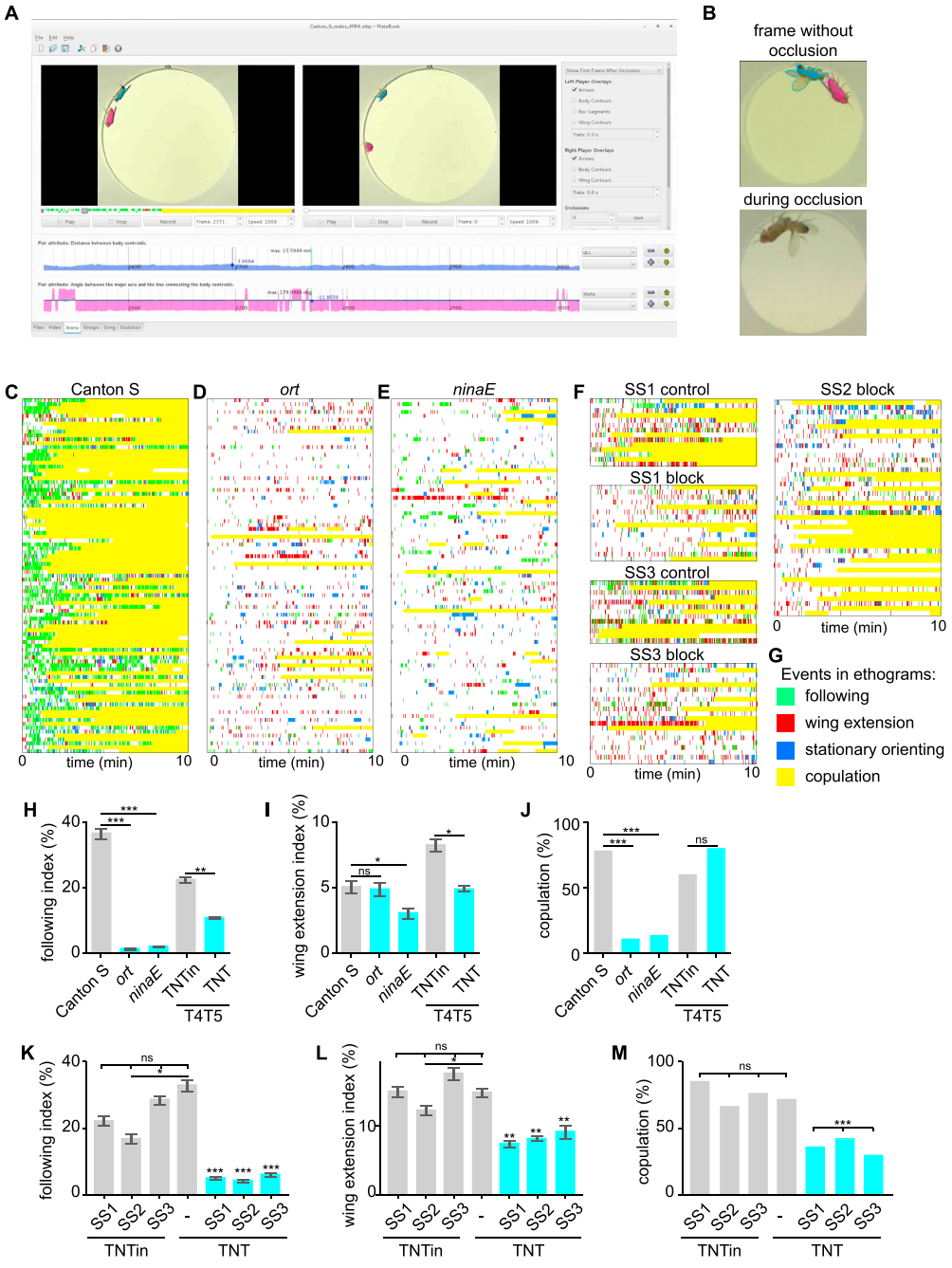


Figure S1. MateBook Software for Automated Analysis of Courtship Behavior, Related to Figures 1 and 2

(A) Snapshot of the graphical user interface (GUI) for MateBook software at the arena tab. This tab displays the current video frame (left, green vertical line in time series at bottom) and the frame immediately after the last occlusion (right), and a time series of two selected attributes for the male (blue) and female (pink; here the distance [dcc] and angle θ [theta] to the other fly).

(B) Examples of non-occluded (left) and occluded (right) video frames. Individual bodies and wings are segmented only in non-occluded frames.

(legend continued on next page)

(C–E) Automatically-generated ethograms for single-pair courtship assays between a male of the indicated genotype and a *Canton S* female. Each line represents a single assay. n = 50 for each genotype.
(F) Ethograms for LC10 block and control males paired with *Canton S* females. SS1 control *(LC10-SS1 > IMPTNTQ)*, n = 13; SS3 control (*LC10-SS3 > IMPTNTQ)*, n = 13; SS2 block, (*LC10-SS2 > TNT)*, n = 46; SS3 block (*LC10-SS3>TNT*), n = 21; SS1 block (*LC10-SS1 > TNT)*, n = 21.
(G) Courtship events detected by MateBook.
(H, I, K, L) Courtship parameters derived by MateBook for the assays shown in (C–F) plus parental control lines, with genotypes indicated. Data are mean ± s.e.m. ***p < 0.001, *p < 0.05, n.s p > 0.05 for comparison with control (*Canton S*, UAS-TNT or UAS-IMPTNT-Q) males, Student's t test with Bonferroni correction.
(J and M) Percentage of males copulating in assays shown in (C-F) plus parental control lines, respectively. ***p < 0.0001 for comparison with Canton S (J) or control (J, M) males, Fisher's exact test with Bonferroni correction.

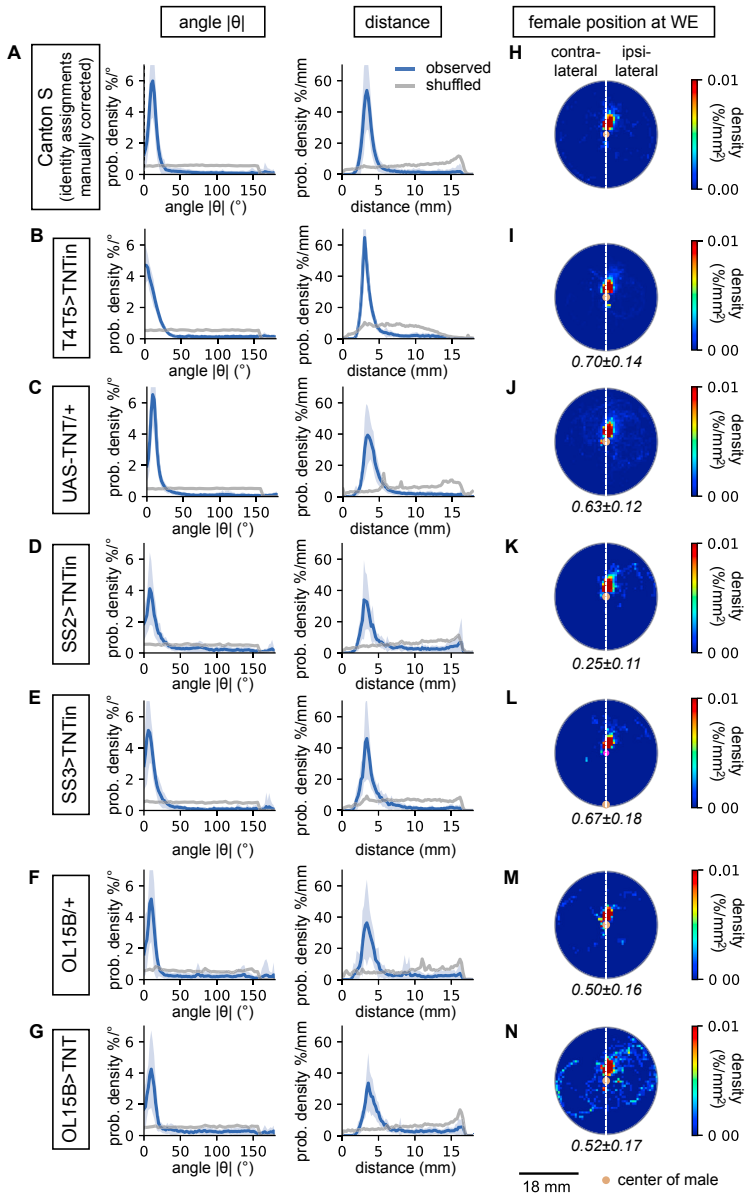


Figure S2. Visual Input Mediates Directed Courtship, Related to Figures 1 and 2

(A–G) Probability density functions of the absolute angle θ (left) and distance (right) for single-pair courtship assay between males of the indicated genotype and Canton S females, for actual (blue) and shuffled (gray) data. Data are mean ± SEM.

(H–N) Heatmap of probability density function for position of female relative to the male ellipse center for every frame in which the male extends a wing.

(A and H) MateBook derived data for *Canton S* flies after manually correcting identity assignments across occlusions, n = 93. Compare to data from the uncorrected videos shown in Figures 1B and 1F.

(B and I) T4T5 control (*SS324>IMPTNT-Q*) males, n = 12.

(C and J) UAS-TNT parental line (*UAS-TNT/+*) males, n = 27.

(D and K) SS2 > TNTin (*LC10-SS2>IMPTNT-Q*) males, n = 40.

(E and L) SS3 > TNTin (*LC10-SS3>IMPTNT-Q*) males, n = 18.

(F and M) OL15B control (*OL15B/+*) males, n = 25.

(G and N) OL15B block (*OL15B>TNT*) males, n = 16.

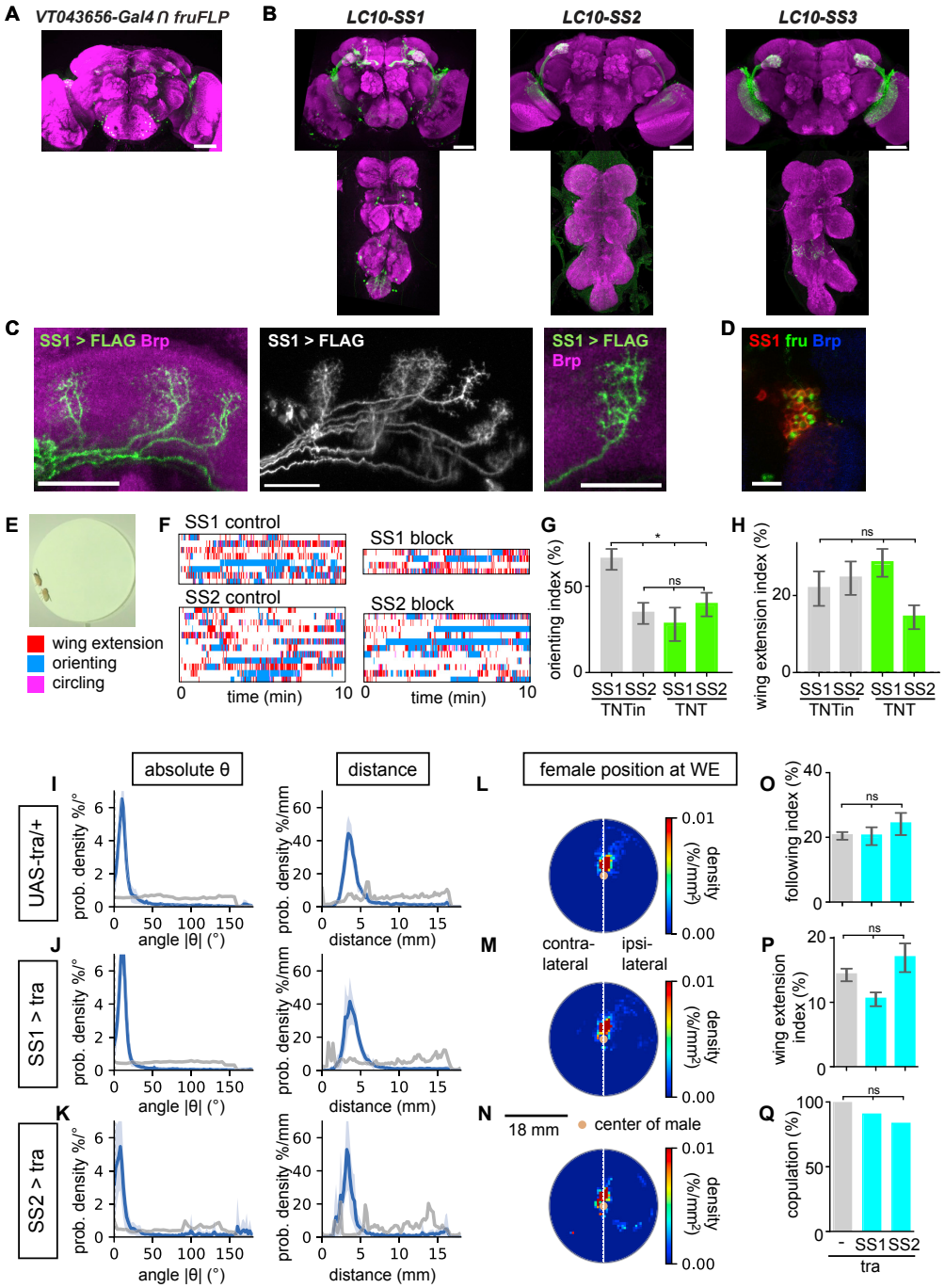


Figure S3. LC10 Neurons Do Not Affect Courtship Arousal and Are Functionally Monomorphic, Related to Figure 2

(A and B) Confocal images of brain and ventral nerve cord of 5-7d-old males from *VT043656-GAL4*, *LC10-SS1*, *LC10-SS2* and *LC10-SS3* > *mCD8:Tomato* labeled with anti-RFP (shown in green) and anti-Brp (magenta) antibodies. Scale bars are 50 μm.

(legend continued on next page)

(C) Single cells in *LC10-SS1* males visualized using the multi-color flip-out technique (Nern et al., 2015), labeled with antibodies against FLAG (green) and Brp (magenta) or FLAG alone (middle panel). Scale bars are 25 μm.
(D) Confocal image of brain of LC10-SS1 > RFP, *fru-LexA* > *lamin:GFP* male labeled with anti-RFP (red) and GFP (green), and counterstained with anti-Brp (blue). 61.13 ± 0.39% (mean ± s.e.m.) of LC10 neurons expressing SS1 are co-labeled with *fru-LexA* in males aged 5-7 days old. Scale bar is 10 μm.
(E) Snapshot of courtship arena showing one test male and a decapitated female.
(F) Ethograms for males of the indicated genotypes paired with decapitated *Canton S* females. SS1 control (*LC10-SS1>IMPTNT-Q*), n = 8; SS1 block (*LC10-SS1>TNT*), n = 4; SS2 control (*LC10-SS2>IMPTNT-Q*), n = 12; SS2 block (*LC10-SS2>TNT*), n = 11.
(G and H) Orienting and wing extension indices derived from courtship assays shown in (F). Data are mean ± SEM.
(I–Q) Courtship data for males with feminized LC10 neurons. Control *UAS-tra* males, n = 16; *LC10 -SS1 > tra* males, n = 22; *LC10 -SS2 > tra* males, n = 12. (I–K) Probability density functions of the absolute angle θ (left) and distance (right) for single-pair courtship assay between males of the indicated genotype and Canton S females, for actual (blue) and shuffled (gray) data. Data are mean ± SEM. (L–N) Heatmap of probability density function for position of female relative to the male ellipse center for every frame in which the male extends a wing. (O–Q) Following (O), wing extension (P) and copulation (Q) indices. (O and P) Data are mean ± SEM.

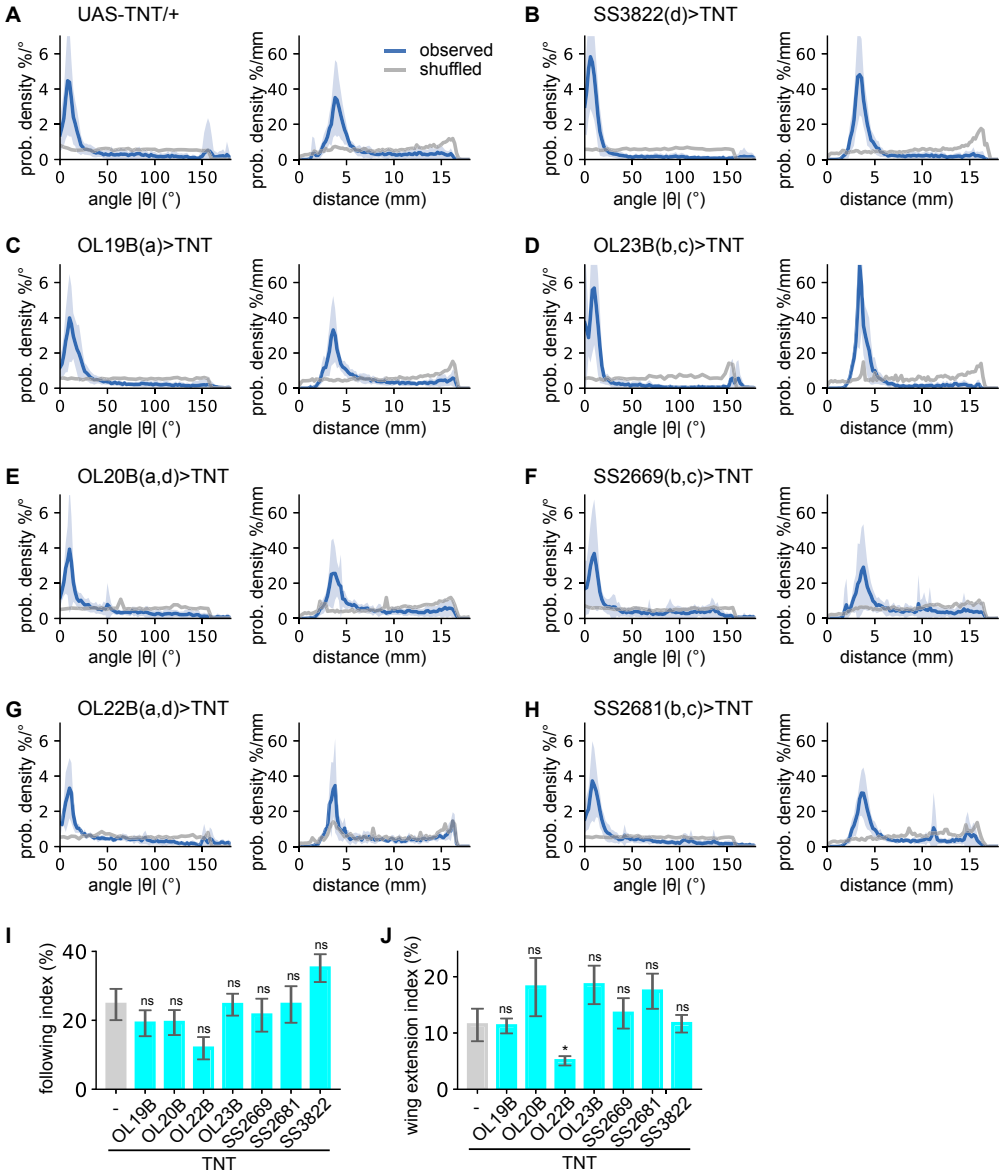


Figure S4. Courtship Behavior for Split-GAL4 Lines Differentially Expressed in LC10 Subtypes, Related to Figure 2

(A–H) Probability density functions for angle θ (left) and distance (right) from male to female, as shown as in Figure 1. Data are mean ± SEM. (A) TNTe control (UAS-TNT/+), n = 16. (B) *SS3822>TNT*, n = 13. (C) *OL19B>TNT*, n = 12. (D) *OL23B>TNT*, n = 6. (E) *OL20B>TNT*, n = 14. (F) *SS2669>TNT*, n = 16. (G) *OL22B>TNT*, n = 4. (H) *SS2681>TNT*, n = 10.

(I and J) MateBook derived following (I) and wing extension (J) indices for males of the indicated genotypes paired with Canton S males. LC10 split-GAL4 are as reported in (Wu et al., 2016). *p < 0.05, n.s p > 0.05 for comparison with control (TNT parental line) males, Student's t test with Bonferroni correction. Data are mean ± SEM.

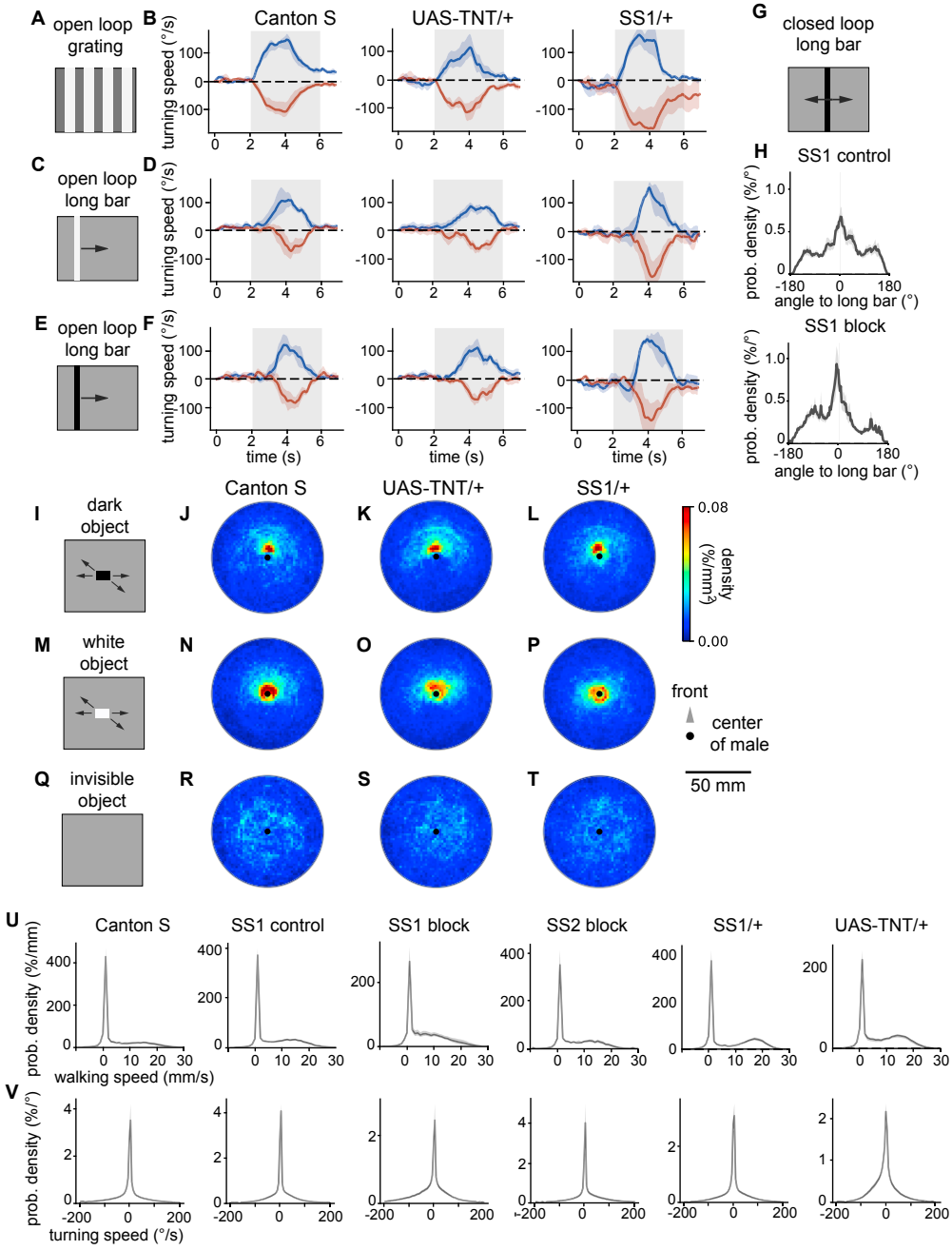


Figure S5. Behavioral Responses of Tethered Males to Visual Stimuli, Related to Figure 3

(A, C, and E) Schematics showing the visual stimuli presented in open loop to the tethered male walking on an air-suspended ball. (A) Vertical gratings with wavelength of 20° moving for 2 s at 20°/s. (C and E) Vertical 10° bright (C) or dark (E) bar sweeping at 60°/s; as in Figure 3.

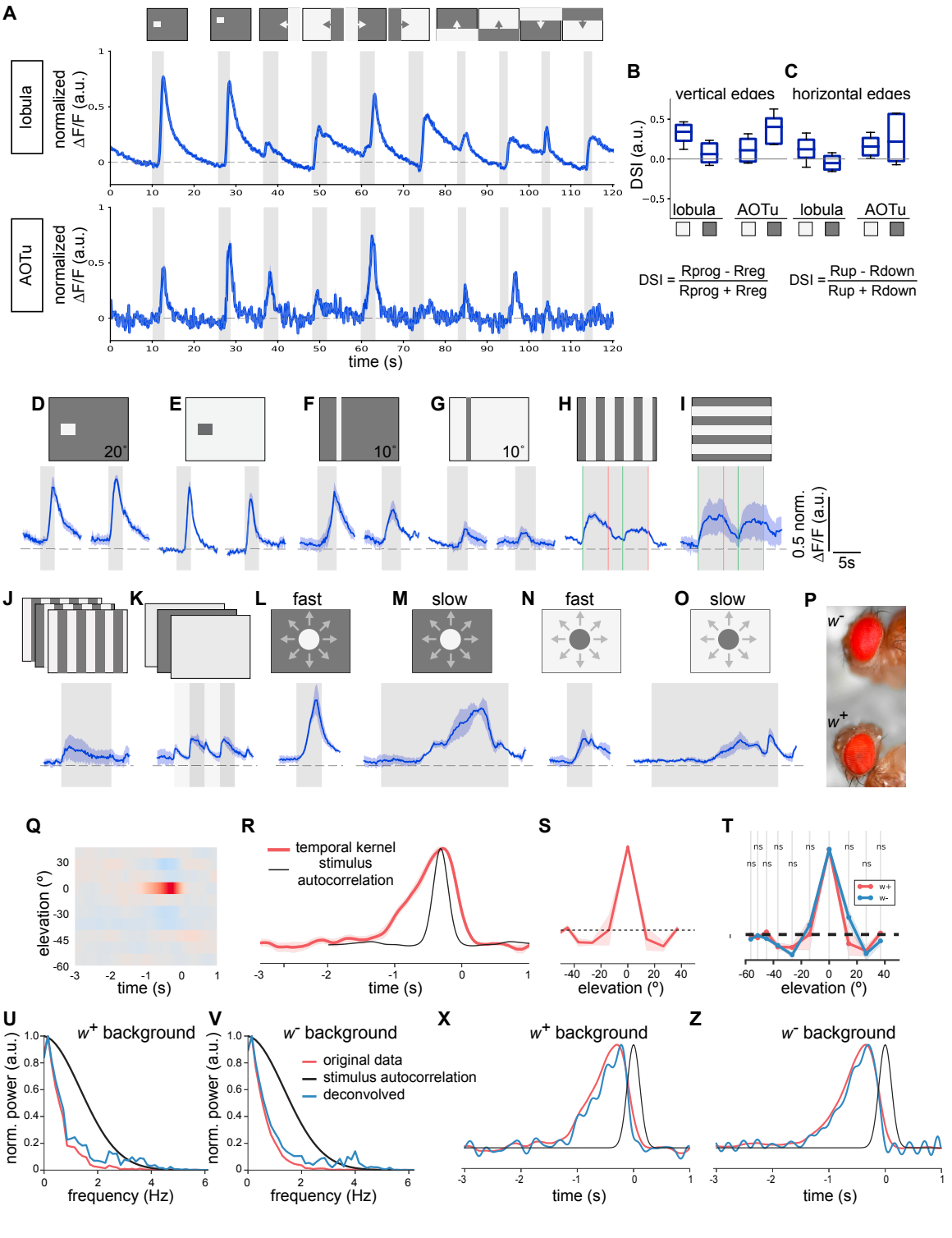
(B, D, and F) Turning speed over time in response to visual stimuli shown in (A, C and E), respectively. Data are mean ± s.e.m. responses to clockwise (blue) or counter-clockwise (red) motion. *Canton S*, n = 7; TNT parental line (*UAS-TNT/+*), n = 9; SS1/+ (*LC10-SS1/+*), n = 6. Data are mean ± SEM.

(G) Schematic of vertical 10° dark bar presented in closed loop.

(H) Probability density functions of bar position relative to the male for SS1 control (*LC10-SS1/+*), n = 9; SS1 block (*LC10-SS1>TNT*), n = 7. Data are mean ± SEM.

(legend continued on next page)

(I, M, and Q) The 3-dimensional virtual object was either dark (I), bright (M), or invisible (Q). The latter controls for the reaction of the male to the virtual cylinder itself. As in Figure 3.
(J–T) Heatmaps of probability density function for the position of the dark (J-L), bright (N-P) or invisible (R-T) object relative to the male center.
(J, N, and R) *Canton S*, n = 7.
(K, O, and S) UAS-TNT/+ (*UAS-TNT/+*), n = 9.
(L, P, and T) SS1/+ (*LC10-SS1/+*), n = 10.
(U and V) Distribution of walking(top) and turning (bottom) speeds of tethered males of indicated genotypes on the air-suspended ball. Data shown here are from same males as (J-T), and the corresponding males in Figures 3I–3S. Data are mean ± SEM.



(legend on next page)

Figure S6. Calcium Responses of LC10 Neurons to ON and OFF Edges in Males, to Test-Stimuli in Females, and to Motion-Noise Stimulus in w^+ Background, Related to Figures 4 and 5

(A) Schematic of stimuli presented during the periods indicated by shading in calcium response traces for LC10 processes in the lobula (middle) and AOTu (bottom; *LC10-SS1>GCaMP6m*). Data are mean ± s.e.m., n = 68 ROIs in 6 flies for lobula, 24 ROIs in 4 flies for AOTu.

(B) Direction selectivity index (DSI) for progressive and regressive motion. Rprog indicates maximum responses to progressive motion; Rreg indicates maximum responses to regressive motion.

(C) Direction selectivity index for upward and downward motion. Rup indicates maximum responses to upward motion; Rdown indicates maximum responses to downward motion.

(D–O) Schematic of visual stimuli (top) and calcium responses in LC10 processes in the lobula of female flies (*LC10-SS1>GCaMP6m*). Stimuli as in Figure 4. Data are mean ± s.e.m., acquired from n = 55 ROIs in 4 females SS1 > GCaMP6m flies.

(P) Eye color in w^+ and w^- background of genotypes used for imaging.

(Q) Correlation of calcium transients observed in LC10 arborizations in w^+ background males with spatiotemporal velocity profiles and the same visual stimuli as in Figures 5C–5F. The mean of aligned spatiotemporal profiles is presented; n = 97 ROIs in 5 flies.

(R) The temporal receptive field and the stimulus autocorrelation as a function of time for data in Q.

(S) The spatial receptive field as a function of relative elevation for data in Q.

(T) Comparison between spatial profiles measured in w^+ (this figure) and w^- (Figures 5G–5I) backgrounds. n.s., $p > 0.05$. (R, S, and T) Data are mean ± SEM.

(U and V) Normalized power spectra of the temporal component of the stimulus-response cross-correlation function (red), with Wiener deconvolution (blue) with the stimulus power spectrum (black) for measurements in w^+ (U) and w^- (V) backgrounds.

(X and Z) The temporal receptive field (blue and red lines) plotted with the stimulus autocorrelation (black), not peak-aligned. The normalized stimulus response cross-correlation (red) was deconvolved with Wiener deconvolution (blue) with the stimulus auto-correlation function (black) for w^+ (H) and w^- (I) backgrounds. The Wiener deconvolution compensates for the blur of the data caused by the stimulus auto-correlation, but not by the calcium indicator (GCaMP6m). This deconvolution results in very similar time course of the estimated temporal filtering properties of LC10 neurons. The noise in the measured calcium transients at higher frequencies is enhanced by the deconvolution producing high frequency ripples in the time domain.

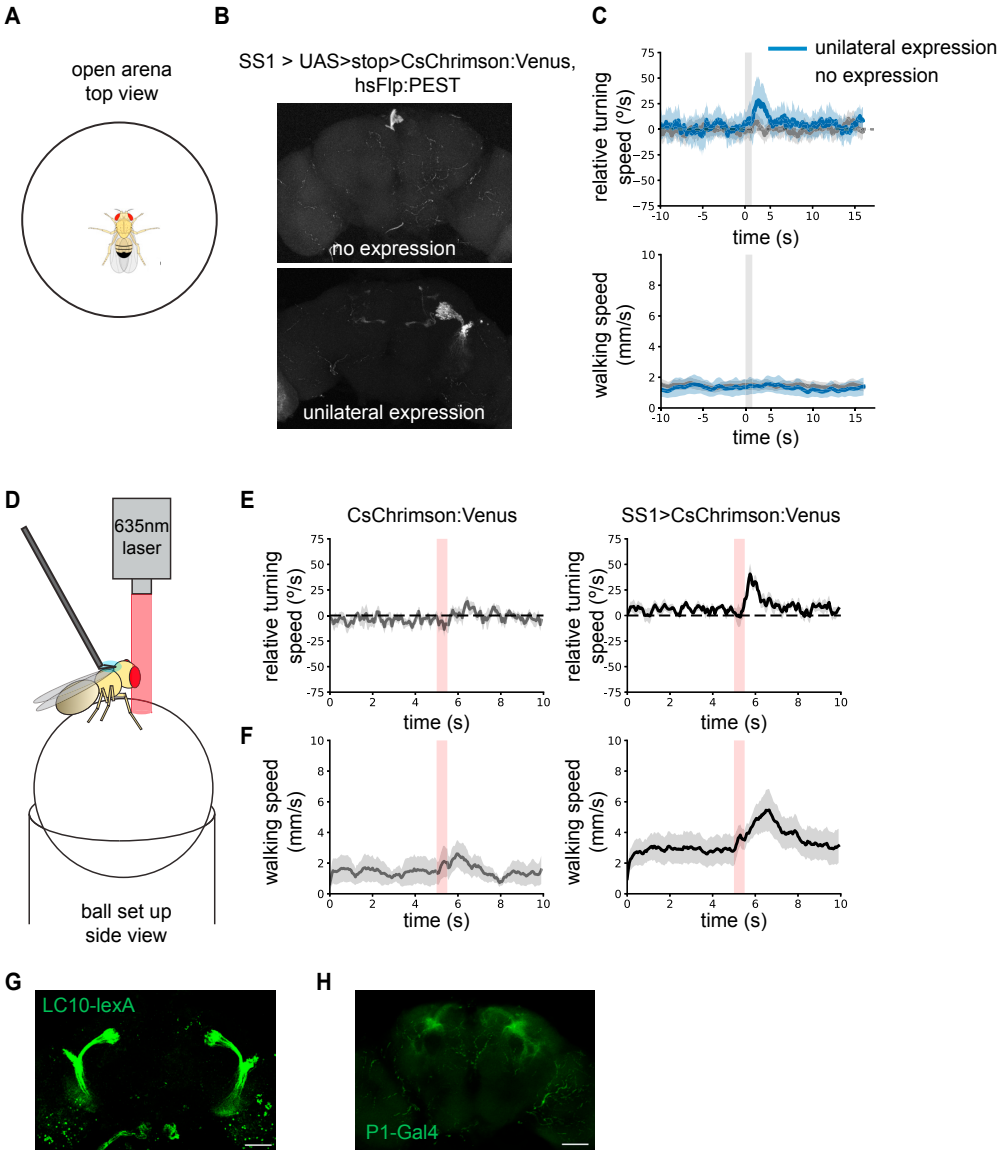


Figure S7. Unilateral Activation of LC10 Elicits Ipsilateral Turning, Related to Figure 7

(A) Schematic of the open arena, in which the red LED light stimulation was provided below the arena.

(B) Stochastic activation of LC10 with *LC10-SS1 > FRT-stop-FRT-CsChrimson:Venus*. Examples of brains with no expression (n = 19) and unilateral expression (n = 4) in LC10 neurons.

(C) Walking and relative turning speeds during a 0.5 s red LED stimulation. Data are mean ± SEM.

(D) Schematic showing unilateral activation performed with spatially restricted laser spot on a tethered male walking on an air suspended ball.

(E and F) Walking and relative turning speeds during 0.5 s exposure to spatially restricted laser spot to one side of the head. Positive relative turning indicates turning toward the stimulated side. Control (*CsChrimson:Venus/+*), n = 8; SS1 > CsChrimson:Venus (*LC10-SS1 > CsChrimson:Venus*), n = 11 males. Data are mean ± SEM.

(G) Confocal image of brain of 5-7d-old males from *LC10-lexA > myr::GFP* labeled with anti-GFP (green). Scale bar is 50μm.

(H) Confocal image of brain of 5-7d-old males from *LC10-lexA, P1-GAL4 > TrpA1:myc* labeled with anti-myc (green). Scale bar is 50μm.

2.3 GLUTAMATE SIGNALING IN THE FLY VISUAL SYSTEM

SUMMARY This study uses the glutamate indicator iGluSnFR for functional characterization of glutamatergic neurons in the fly motion vision circuitry.

In Manuscript 1, we used the calcium indicator GCaMP6f to map response properties of input neurons to T4 and T5 cell. However, it is known that this calcium indicator possesses relatively slow kinetics with a decay time constant around 200–400ms. The glutamate sensor iGluSnFR has a much shorter decay time constant and therefore represents an alternative to calcium imaging for glutamatergic neurons. In this study, we first showed using immunolabeling techniques that the vesicular glutamate transporter VGlut co-localizes with axon terminals of L1, Mi9 and LPi neurons. We then expressed the fast glutamate indicator iGluSnFR in these cell types and confirmed that this glutamate sensor can be used for two-photon imaging in *Drosophila*. Using glutamate signals as a read-out, we mapped spatiotemporal receptive fields of L1 and Mi9 cells. Our results matched previous descriptions of the functional properties of these neurons. Importantly, we could obtain a more precise quantification of the low-pass filter time constant of Mi9 cells, which we had only estimated by deconvolution of calcium signals in Manuscript 1. Finally, we could gather evidence that LPi neurons are also direction selective in their glutamate responses. Overall, we described the response dynamics of L1, Mi9 and LPi neurons with much higher temporal precision than before and found that iGluSnFR is a suitable tool for functional imaging in *Drosophila*.

This article was published in *iScience* in September 2018 (Richter et al., 2018).

AUTHORS Florian G. Richter, Sandra Fendl (co-first author), Jürgen Haag, **Michael S. Drews** & Alexander Borst

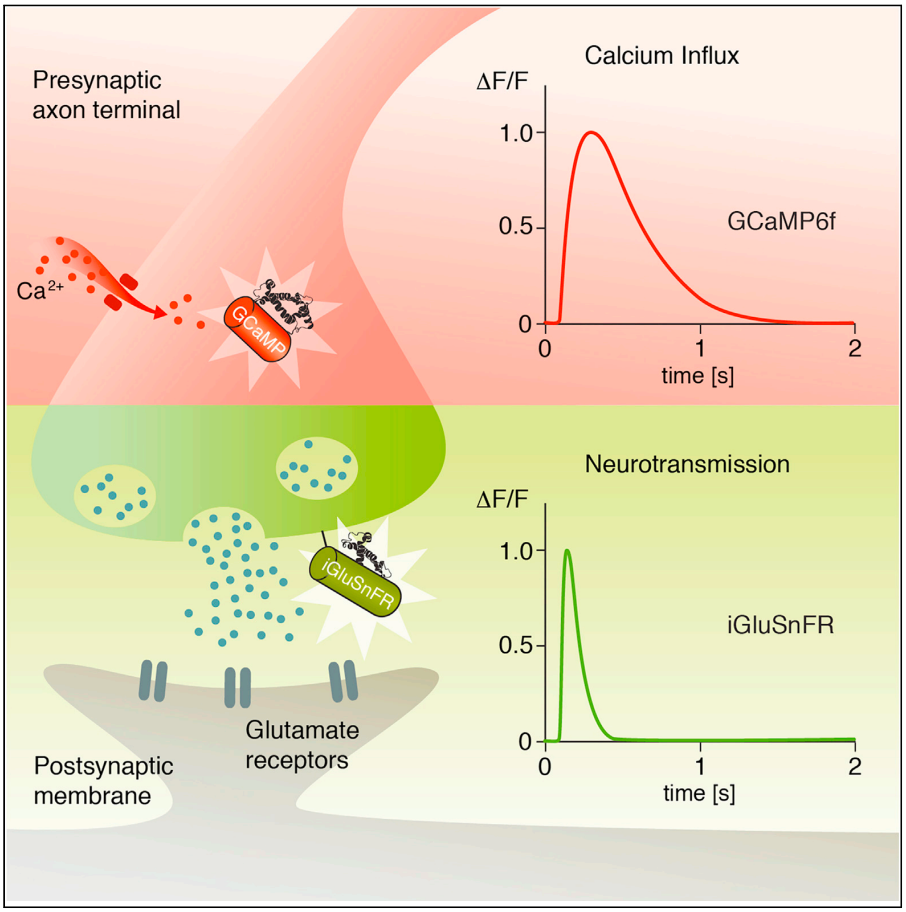
CONTRIBUTIONS F.G.R., S.F., and A.B. conceived the study and designed the experiments. F.G.R. conducted and analyzed the imaging experiments for Mi9 and L1. S.F. performed and analyzed all stainings. J.H. performed and analyzed the LPi experiments. **M.S.D.** performed data analysis and model fitting of the receptive fields. F.G.R. wrote the manuscript with the help of all authors.

iScience

CellPress

Article

Glutamate Signaling in the Fly Visual System



Florian G. Richter, Sandra Fendl, Jürgen Haag, Michael S. Drews, Alexander Borst

frichter@neuro.mpg.de (F.G.R.)
aborst@neuro.mpg.de (A.B.)

HIGHLIGHTS

The glutamate sensor iGluSnFR is suitable for 2-photon imaging in the fruit fly

Response properties obtained with iGluSnFR are much faster than those with GCaMP6f

Spatial aspects of receptive fields are preserved between indicators

Richter et al., iScience 7, 85–95
September 28, 2018
https://doi.org/10.1016/j.isci.2018.08.019

iScience

CellPress

Article

Glutamate Signaling in the Fly Visual System

Florian G. Richter,[1,2,*] Sandra Fendl,[1,2] Jürgen Haag,[1] Michael S. Drews,[1] and Alexander Borst[1,3,*]

SUMMARY

For a proper understanding of neural circuit function, it is important to know which signals neurons relay to their downstream partners. Calcium imaging with genetically encoded calcium sensors like GCaMP has become the default approach for mapping these responses. How well such measurements represent the true neurotransmitter output of any given cell, however, remains unclear. Here, we demonstrate the viability of the glutamate sensor iGluSnFR for 2-photon *in vivo* imaging in *Drosophila melanogaster* and prove its usefulness for estimating spatiotemporal receptive fields in the visual system. We compare the results obtained with iGluSnFR with the ones obtained with GCaMP6f and find that the spatial aspects of the receptive fields are preserved between indicators. In the temporal domain, however, measurements obtained with iGluSnFR reveal the underlying response properties to be much faster than those acquired with GCaMP6f. Our approach thus offers a more accurate description of glutamatergic neurons in the fruit fly.

INTRODUCTION

To understand how neural circuits operate and carry out certain computations, it is essential to observe the signals that are transmitted from cell to cell. Synaptic transmission via chemical synapses proceeds in four major stages: (1) Depolarization in the presynapse opens voltage-gated calcium channels. (2) The resulting calcium influx leads to the fusion of transmitter-filled vesicles and the presynaptic membrane. (3) Transmitter molecules are released into the synaptic cleft where they diffuse and bind receptors in the postsynaptic membrane. (4) The subsequent activation of these receptors leads to opening or closing of ion channels, either directly or indirectly, with the resulting ion flux ultimately changing the postsynaptic membrane conductance and potential (reviewed in [Di Maio, 2008]). This fundamental signaling cascade, from electric potential to calcium to transmitter release to postsynaptic electric potential, orchestrates computation within any neuronal circuit.

For monitoring voltage changes, electrophysiology is the default approach. Here, direct observations of both de- and hyperpolarization in pre- or postsynaptic cells are possible. Due to the position or size of many neurons, however, direct single-cell recordings are often not feasible and have to be replaced by indirect extracellular recordings or optical imaging. Only recently genetically encoded voltage indicators (GEVIs) have emerged as powerful tools for recording neuronal activity (Cao et al., 2013; Jin et al., 2012; St-Pierre et al., 2014; Tsutsui et al., 2013; Yang et al., 2016). Experiments with optical voltage indicators such as ASAP2f that are compatible with 2-photon imaging, however, remain challenging due to weak signal-to-noise ratio (Yang et al., 2016). The fluorescence level of genetically encoded calcium indicators (GECIs) is thought to correlate with transmitter release and is therefore suitable for identifying the crucial signal to the postsynaptic cell (Zucker, 1993). Although GECIs are being improved continuously and some variants were designed to have especially fast kinetics (e.g., GCaMP6f [Chen et al., 2013]), temporal resolution is still limited due to calcium buffering (Borst and Abarbanel, 2007). This usually leads to decay constants in the order of several hundreds of milliseconds that vary depending on the system under observation (Arenz et al., 2017; Chen et al., 2013). For glutamatergic neurons, a tool to potentially overcome these limitations is the recently developed fast glutamate sensor iGluSnFR (Marvin et al., 2013).

Visual motion detection is a canonical example for computation in neural microcircuits. Prevalent models posit that, in both mammalian retina and fly visual system, local direction selectivity emerges from the nonlinear interaction between precisely tuned spatiotemporal filters (Barlow and Levick, 1965; Von Hassenstein and Reichardt, 1956). Recent work in connectomics on the visual system of *Drosophila melanogaster* has revealed this computation to be implemented by a circuit that consists of only a few dozen individual cells (Takemura et al., 2017). The optic lobe is the largest neuropil in the fruit fly's brain and consists of the

[1]Max-Planck-Institute of Neurobiology, 82152 Martinsried, Germany

[2]These authors contributed equally

[3]Lead Contact

*Correspondence: frichter@neuro.mpg.de (F.G.R.), aborst@neuro.mpg.de (A.B.)

https://doi.org/10.1016/j.isci.2018.08.019

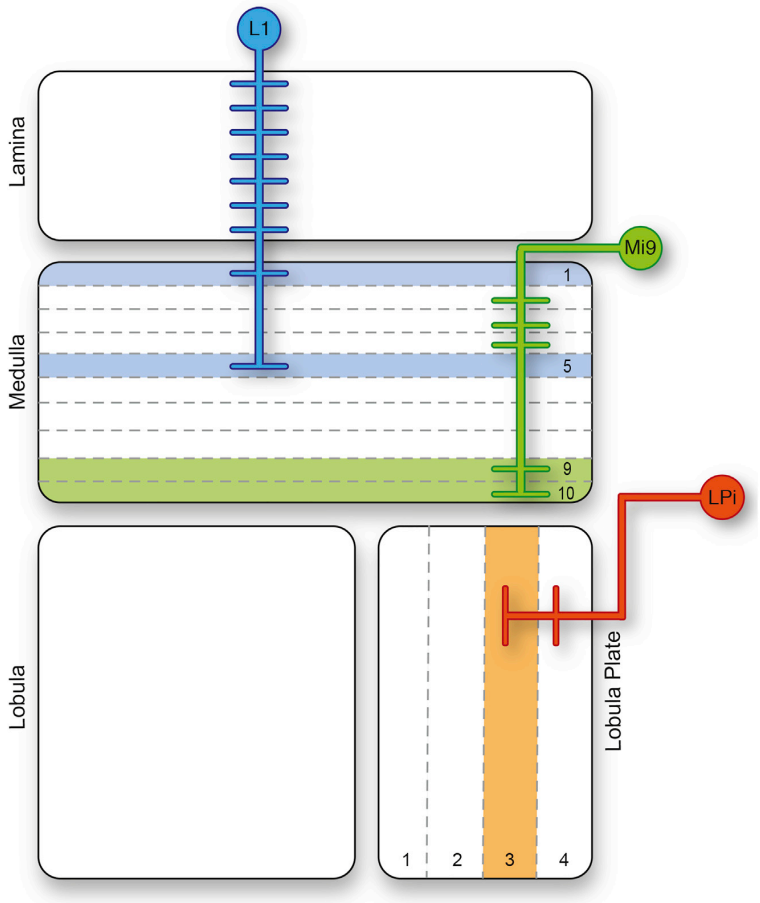


Figure 1. Schematic of the *Drosophila* Optic Lobe

Schematic of the *Drosophila* optic lobe with glutamatergic cell types in the motion vision pathway. The three cell types are not directly connected to each other but play an import role in the circuit. For the sake of simplicity, postsynaptic partners of the glutamatergic neurons are not displayed but can be reviewed in Mauss et al. (2015) and Takemura et al. (2011, 2017). Colored layers indicate area where we imaged glutamate release of the respective cell type.

four consecutive neuropils: lamina, medulla, lobula, and lobula plate (Figure 1). Lamina monopolar cells L1 and L2, among others, receive direct photoreceptor input and feed into two parallel pathways (Bausenwein et al., 1992; Bausenwein and Fischbach, 1992; Borst, 2014; Clark et al., 2011; Joesch et al., 2010; Rister et al., 2007; Shinomiya et al., 2014; Silies et al., 2013; Takemura et al., 2017; Tuthill et al., 2013). The ON pathway processes the motion of light increments, whereas the OFF pathway processes the motion of light decrements only (Eichner et al., 2011; Joesch et al., 2013, 2010). Among the medulla interneurons that connect the lamina cells to direction-selective T4 and T5 neurons (Maisak et al., 2013; Takemura et al., 2017), we find the glutamatergic cell Mi9 that has been characterized with a receptive field responsive to OFF in the center and an antagonistic ON surround (Arenz et al., 2017; Strother et al., 2017). T4 and T5 neurons each come in four subtypes, tuned to one of the four cardinal directions, and project, according to their preferred direction, to one of the four layers in the lobula plate. Here, T4 and T5 cells make excitatory cholinergic connections onto the dendrites of large tangential cells as well as onto inhibitory lobula plate interneurons (LPis). These neurons in turn inhibit large field tangential cells in the adjacent layer during null direction motion and thus increase their flow-field selectivity (Hausen et al., 1980; Hopp et al., 2014; Schnell et al., 2010; Scott et al., 2002; Wasserman et al., 2015). To provide this inhibition, LPis release glutamate onto the glutamate

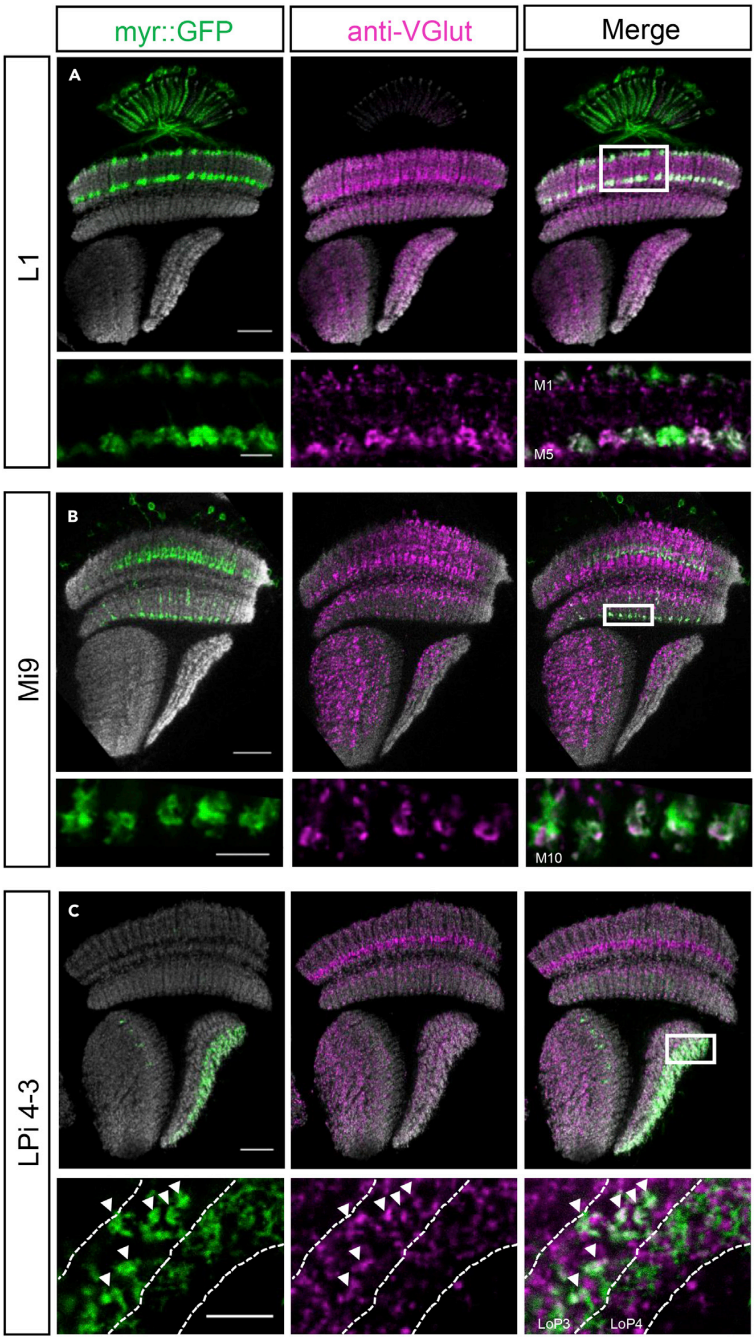


Figure 2. Vesicular Glutamate Transporter VGlut Localizes to Axon Terminals of L1, Mi9, and LPi4-3 Neurons Indicating their Glutamatergic Phenotype

(A–C) Upper rows show overviews of optic lobes with L1 (A), Mi9 (B), and LPi4-3 (C) labeled with myr::GFP (green), background staining against bruchpilot brp (gray), and anti-VGlut staining (magenta). In the lower rows higher magnifications of axon terminals of L1, Mi9, and LPi4-3 neurons are depicted (sections marked with white boxes in overview images).

Figure 2. *Continued*
(A) L1 axon terminals in medulla layers 1 and 5 show overlapping signal with anti-VGlut staining.
(B) VGlut protein co-localizes with Mi9 axons in layer 10 of the medulla.
(C) Lobula plate intrinsic neurons LPi4-3 have their dendrites in layer 4 and project their terminals to layer 3. Labeled with arrowheads are LPi boutons in layer 3 showing overlapping signal with anti-VGlut staining. Shown here are single planes of confocal stacks. Scale bar for overview of optic lobes is 20 μm. For higher magnification close-ups the scale is 5 μm. White dashed lines in the lower panel are manually drawn and indicate layers of the lobula plate.

receptor GluClα, which is an inhibitory glutamate receptor only found in invertebrates (Liu and Wilson, 2013; Mauss et al., 2015, 2014).

The exact biophysical mechanisms by which T4 and T5 become direction selective remain unclear. To understand on a cell-by-cell level how direction selectivity is achieved, precise measurements of the signals transmitted between neurons are crucial. In this study, we focus on the final stage of the synaptic signaling cascade, i.e., transmitter release. First, we confirm the neurotransmitter phenotype of all known glutamatergic cell types (L1, Mi9, LPi) in the *Drosophila* motion vision pathway. Second, using the recently developed fast glutamate sensor iGluSnFR (Marvin et al., 2013), we comprehensively characterize their spatiotemporal response profiles and compare them with the ones obtained expressing the genetically encoded calcium indicator GCaMP6f (Chen et al., 2013).

RESULTS

The Vesicular Glutamate Transporter VGlut Localizes to Axon Terminals of L1, Mi9, and LPi4-3 Neurons

VGlut or DVGLUT (CG9887) is the only vesicular glutamate transporter known in *Drosophila*. VGlut is located in the vesicle membrane of glutamatergic neurons where it fills the synaptic vesicles with glutamate. The protein localizes to presynaptic terminals of all known glutamatergic neuromuscular junctions (NMJs) as well as to synapses throughout the CNS neuropil in *Drosophila* (Daniels, 2004). Hence, VGlut is the most commonly used marker for glutamatergic neurons. Several antibodies have been raised against VGlut to identify glutamatergic neurons in the nervous system of the fruit fly (Daniels, 2004; Mahr and Aberle, 2006).

Recent studies revealed the glutamatergic phenotype of L1, Mi9, and LPi neurons—each of them a crucial element of the motion vision pathway of the fruit fly (Joesch et al., 2010; Kolodziejczyk et al., 2008; Mauss et al., 2015; Takemura et al., 2017, 2011). The somata of these cell types showed positive immunoreactivity against the VGlut antibody, which was raised against a C-terminal peptide—CQMPSYDPQGYQQQ (Daniels, 2004). Interestingly, this antibody labeled mainly cell bodies of designated neurons. Since it is known that the vesicular glutamate transporter VGlut is localized to axon terminals, we investigated the glutamatergic transmitter phenotype of L1, Mi9, and LPi4-3 in more detail. We used a different anti-VGlut antibody (Mahr and Aberle, 2006), which only labels neuronal arborizations in the optic lobe neuropil and no somata. In general, the VGlut protein is highly abundant throughout all four neuropils of the optic lobe (Figure 2).

The axon terminals of L1 neurons show clear overlap with the anti-VGlut signal in layer M1 and M5 of the medulla (Figure 2A). The vesicular glutamate transporter VGlut resides at the presynaptic sites of L1 neurons, which indicates their glutamatergic phenotype. In layer M10 of the medulla, the same is found for Mi9 neurons: VGlut staining in this layer is co-localized with GFP-labeled Mi9 axon terminals (Figure 2B). This suggests that Mi9 neurons are glutamatergic and that they are the only source of glutamate in layer M10 of the medulla. Furthermore, we found an overlapping signal of LPi4-3 terminals in layer 3 of the lobula plate and anti-VGlut staining (Figure 2C). This confirms recent findings (Mauss et al., 2015) that described LPi neurons as glutamatergic, being presynaptic only in one of the two layers where it arborizes.

In summary, we could show that the protein VGlut localizes to axon terminals of the glutamatergic neurons L1, Mi9, and LPi4-3.

Faster Sensor Kinetics Enable More Precise Characterization of Visual Interneurons

One commonly used approach to characterize a sensory neuron is to find its preferred stimulus. This can be achieved by using a white noise input and cross-correlating the resulting output with the input (Dayan and

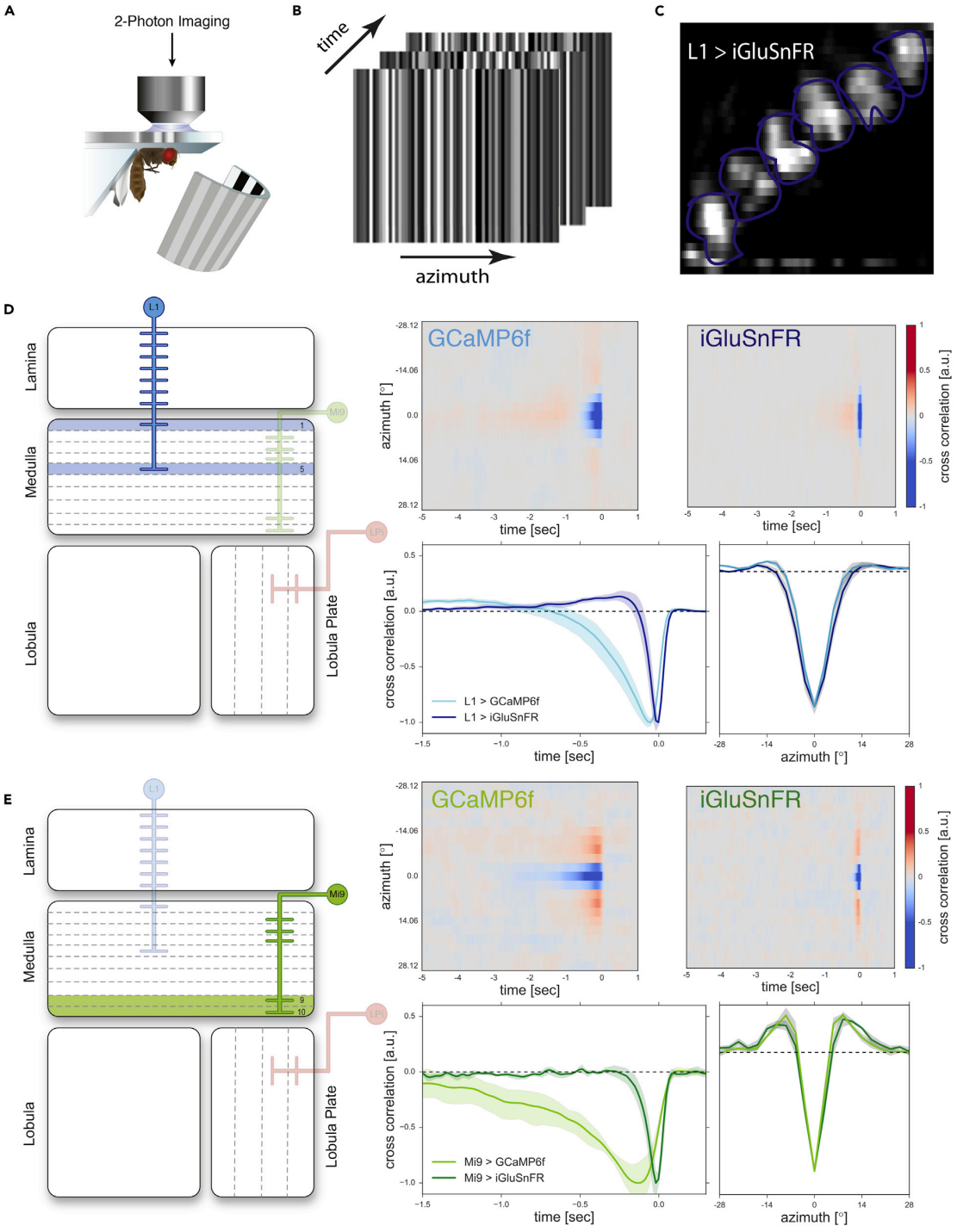

A
2-Photon Imaging
B
time
azimuth
C
L1 > iGluSnFR
D
Lamina
Medulla
Lobula
Lobula Plate
L1
Mi9
LPi
GCaMP6f
iGluSnFR
azimuth [°]
time [sec]
cross correlation [a.u.]
L1 > GCaMP6f
L1 > iGluSnFR
E
Mi9 > GCaMP6f
Mi9 > iGluSnFR

Figure 3. Response Properties of the ON Pathway Columnar Elements L1 and Mi9

(A) Experimental setup: Fly tethered to a plastic holder under the 2-photon microscope looking onto the stimulus arena (see also Transparent Methods).

(B) Schematic of three frames of the white noise stimulus consisting of 64 horizontal bars.

(C) Example of 2-photon image of L1 expressing iGluSnFR. In purple are manually drawn region of interest ROIs.

(D) Left: Schematic of the *Drosophila* optic lobe. The cell type related to the right panel is highlighted. Right upper panel: Averaged aligned spatiotemporal receptive fields after reverse correlation of L1 expressing either the glutamate indicator iGluSnFR (5 flies and 66 cells) or GCaMP6f (5 flies and 60 cells). Cross sections along space and time axes result in receptive fields in right lower panel. Spatial receptive fields do not differ significantly for both indicators. Temporal kernels differ substantially. Impulse responses are shorter for iGluSnFR than for GCaMP6f. Shaded areas indicate a confidence interval of 95%.

(E) Same as (D) only for Mi9 (with iGluSnFR: 5 flies, 26 cells; with GCaMP6f: 5 flies, 50 cells).

Abbott, 2013; French, 1976; Ringach and Shapley, 2004), which yields the linear spatiotemporal receptive field as a result (e.g., Figures 3D and 3E, upper panel). The receptive field of a neuron is defined as the location of a stimulus in space and the time relative to its occurrence in which the neuron's response is modulated by the stimulus. The receptive field also describes the specific filtering properties of a system, in space as well as in time. Here, we use simple first-order low-pass, high-pass, or band-pass filters to quantify these filtering properties using the measured receptive fields. A low-pass filter only allows low frequencies to pass and attenuates high frequencies. Conversely, a high-pass filter attenuates low frequencies and allows high frequencies to pass. A band-pass filter is a combination of a high-pass and a low-pass filter in series, allowing signals within a certain frequency band to pass and attenuating all others (Cruse, 1996). In a linear system, the filters characterized this way are equivalent to the neurons' impulse responses. The temporal impulse response reveals critical aspects of the cellular response kinetics (Dayan and Abbott, 2013; Ringach and Shapley, 2004).

For this reason, we characterized the spatial extent of the receptive fields as well as the response dynamics of all known glutamatergic cells in the motion vision circuit of *Drosophila* L1, Mi9, and LPi4-3. Expressing either the fast version of the genetically encoded calcium indicator GCaMP6f (Chen et al., 2013) or the fast glutamate-sensing reporter iGluSnFR (Marvin et al., 2013) with cell-type-specific Gal4 driver lines, we imaged glutamate and calcium signals in single axon terminals (Figure 3C). To precisely map the receptive fields of these cells, we used a one-dimensional white noise stimulus consisting of 2.8° wide vertical bars covering the full extent of the arena (180°, Figure 3B, see also Methods). The spatiotemporal receptive fields were then determined from the neuron's calcium or glutamate response by reverse correlation. Cross sections through the peak of the spatiotemporal receptive fields along the space axis therefore yield the one-dimensional spatial receptive fields depicted in Figures 3D and 3E. Cross sections along the time axis yield the temporal filtering properties of the neuron (Chichilnisky, 2001; Dayan and Abbott, 2013; French, 1976; Ringach, 2004).

To calculate the spatial extent of the cells' receptive field, we fitted a Mexican hat function (also called difference of Gaussians) that best resembled the center-surround structure of the estimated spatial receptive fields. Both neurons show a small confined center of ~7° for Mi9 and 9–11° for L1. The full width at half maximum of the surround is about 40–50° for L1 and 20–30° for Mi9. Considering the uncertainty of the fitted model parameters, these values are similar and lie in the same order of magnitude when comparing results from imaging with both sensors. In addition, testing the raw data of both conditions against each other we find no significant difference (see Figures S2A and S2B, p value > 0.5, Welch's t test) of spatial receptive fields neither for L1 nor for Mi9. Both neurons show a small confined center of ~7° for Mi9 and 9–11° for L1. The size of the surround has the same order of magnitude for both sensors, 40–50° for L1 and 20–30° for Mi9. This is within the range of uncertainty that the fit is subject to. Testing the raw data of both conditions against each other for the two cell types, however, does not yield a significant difference (see Figures S2A and S2B, right panel).

For a reliable estimation of the time constants of the temporal responses, we transferred the impulse responses of L1 and Mi9 into frequency space and fitted either a first-order low-pass or a first-order band-pass filter to the neurons' responses (see Figures S1C and S1D). For L1, we find that the data are best represented by a band-pass filter. The filter derived from the iGluSnFR signal has a low-pass time constant of 70 ms and a high-pass time constant of about 400 ms (see Figure S1A). The time constants derived from the GCaMP6f signal are significantly larger with low-pass and high-pass time constants of 350 and about 1,180 ms, respectively. For Mi9, we find that the temporal properties are best described by a low-pass filter. The estimated time constant of the Mi9 temporal kernel (Figure 3D, lower

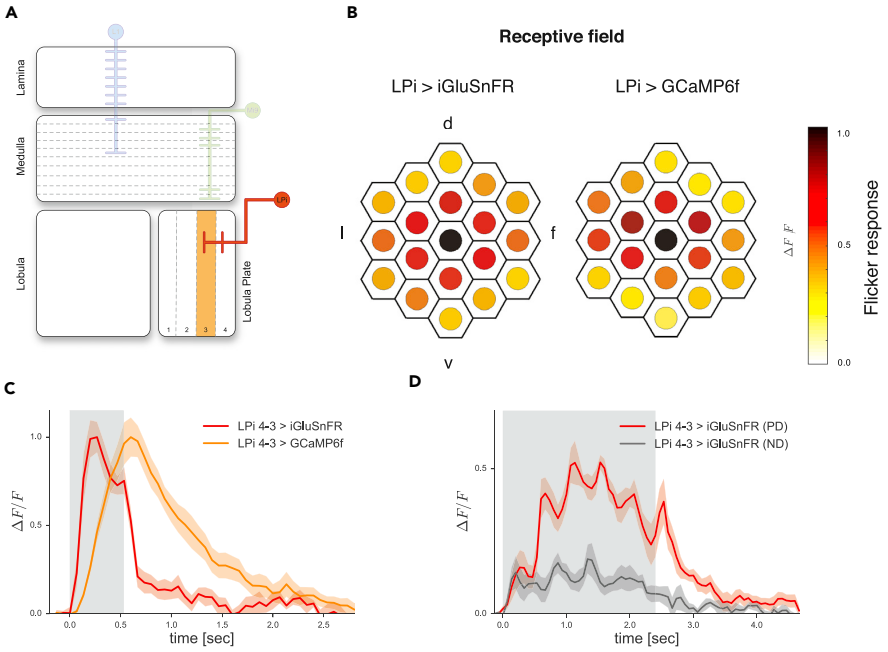


Figure 4. Response Properties of the Direction Selective Lobula Plate Interneuron LPi4-3

(A) Schematic of the *Drosophila* optic lobe with LPi4-3 highlighted.

(B) Comparison of spatial receptive field size of LPi4-3 cells recorded with iGluSnFR (left, n = 24 cells from 7 flies) or GCaMP6f (right, n = 14 cells from 5 flies). The responses of individual cells to flicker stimuli presented at 19 different columnar positions were averaged after alignment to the maximum (in black) and normalization. d, Dorsal; v, ventral; l, lateral; f, frontal.

(C) Time course of LPi4-3 response upon local flicker stimulation. The decay of the signal is faster for iGluSnFR response.

(D) LPi4-3 expressing iGluSnFR show glutamatergic direction selective responses (n = 8 cells from 5 flies). Five consecutive flicker stimuli were shown along the preferred (downward) or null (upward) direction of the neuron, acting as apparent motion. Shaded areas indicate mean ± SEM.

left) is 75 ms when measured with iGluSnFR compared with about 610 ms when measured with GCaMP6f (see Figure S1B).

For both cell types, the temporal kernel of the calcium response can be derived by low-pass filtering the faster glutamate signal. This is because the kinetics of the calcium sensor can be approximated by a low-pass filter when the intracellular calcium concentration is small compared to the KD value of the indicator (Borst and Abarbanel, 2007). For both cells, i.e., L1 and Mi9, we can fit the glutamatergic signal to the calcium signal by filtering it with a low-pass filter with a time constant of 360 ms (see Figures S2A and S2B, left panel). LPis, as motion-selective neurons, are not suitable for white noise analysis. To characterize the response properties of the LPi4-3 (Figure 4A), we first stimulated single ommatidia with local flicker stimuli that were placed precisely onto the lattice of the fly's eye via a custom-built telescopic device (see Transparent Methods and [Haag et al., 2017, 2016]). LPi4-3 cells responded to the individual pulses with different amplitudes, depending on the position of the stimulus (Figure 4C). The maximum response (Figure 4B, black center) of a recorded neuron was then set as the receptive field's center. All other responses to adjacent stimulation are normalized accordingly. Single flicker stimulations in the center of the receptive field show different time courses (Figure 4C) when using the two different indicators. The onset of the calcium response is much slower when compared with the glutamate response. In fact, whereas the glutamate signal shows a short transient peak response and then plateaus after ~500ms, the calcium signal does not resolve any similar details in the time course of the response. The calcium signal decays back to zero in approximately 2 s after stimulus offset, whereas the glutamatergic signals are back at the baseline level in less than 200 ms. This loss-of-response features can be explained by the characteristics of the

calcium indicator, which acts as a low-pass filter (Borst and Abarbanel, 2007). Low-pass filtering the glutamate response (τ = 446 ms, Figure S2C) results in a similar slope and decay as the calcium response. We also asked if the glutamatergic signal of the LPis is indeed direction selective as expected from Mauss et al. (2015). To asses this question we tested LPi4-3 cells with five light pulses of 472 ms duration positioned along the dorsoventral axis of the eye. When stimulated sequentially from dorsal to ventral (Figure 4D), the cell responded more strongly (PD, red line) than when we showed the same stimulus in the opposite direction (ND, black line, paired sample t test, p value < 0.01). We therefore conclude that the sensor is indeed also suitable for resolving glutamatergic direction-selective signals.

DISCUSSION

In this study we showed that all three investigated cell types (L1, Mi9, LPi4-3) express the vesicular transporter for glutamate, VGlut, in their axon terminals (Figure 2). To our knowledge, L1, Mi9, and LPi are the only glutamatergic cells in the *Drosophila* motion vision circuit. Two studies using either antibody stainings (Kolodziejczyk et al., 2008) a Flp-out analysis of the dvGlutCNSIII-Gal4 driver line (heat-shock inducible flipase excises stop-cassette upstream of mCD8-GFP to label only a few cells) (Raghu and Borst, 2011) found L2 cells to be glutamatergic. However, a recent RNA sequencing study that characterized gene expression patterns of more than 60 different cell types of the optic lobe could not confirm the expression of VGlut in L2 (Davis et al., 2018). Although they could identify other cell types like Dm cells, Lai, PB_1, Tm29, and TmY5a as glutamatergic due to their expression of VGlut, none of the other cells in the motion vision circuit (besides L1, Mi9, and LPi) seem to express VGlut. The role of Dm, Lai, PB, Tm29, and TmY5a cells in general and their potential contribution to motion vision in the fly brain are not known to date.

We also demonstrated that the spatial receptive fields measured with the glutamate sensor iGluSnFR are almost identical to the ones measured with the calcium sensor GCaMP6f (Figures 3 and 4). Both neurons possess a local OFF center receptive field with a differently strong antagonistic ON surround. Surround inhibition is a phenomenon frequently found in the early processing stages in visual systems: Bipolar and ganglion cells of the mammalian retina possess receptive fields with an antagonistic center-surround structure (reviewed in Shapley and Lennie, 1985), and first-order interneurons of the insect compound eye share this feature as well (Srinivasan et al., 1982). Functionally, a neuron with a center-surround antagonism acts as a spatial band-pass filter, enhancing the neuron's responses to edges over full field illuminations. Such band-pass filtering reduces redundancy in natural images (Srinivasan et al., 1982). We find such spatial band-pass characteristics for both cell types, L1 and Mi9. Based on their spatial receptive fields, we predict, for instance, no response of Mi9 to wide field dark flashes since the integral of the spatial receptive field is close to zero.

In the time domain, however, the glutamate signal turned out to be much faster than the calcium signal derived from the same cells. Due to their small size, many visual interneurons in the fly brain are inaccessible to electrophysiological recordings, so only a few direct recordings have been reported (Behnia et al., 2014; Gruntman et al., 2018; Juusola et al., 2016). Since data from voltage recordings from L1, Mi9, and LPi are not available so far, a direct comparison with the time constant estimated here is not possible. Simulation studies predicted time constants between 50 and 100 ms for the delayed input to the fly motion-detecting neurons (Eichner et al., 2011; Leonhardt et al., 2016). Since Mi9 is thought to provide this signal to T4 cells, the elementary motion-sensing neurons in the ON pathway, the low-pass time constant of 75 ms estimated here matches this prediction well. In addition, a previous study determined the low-pass time constant for Mi9 to be around 550 ms from calcium imaging experiments. A deconvolution of the filter with an estimated GCaMP kernel led to a resulting time constant of 63 ms (Arenz et al., 2017). This result again is in line with the time constants of the Mi9-iGluSnFR of 75 ms reported here.

In the mammalian CNS, glutamate is the most abundant and major excitatory transmitter (Meldrum, 2000; Traynelis et al., 2010). Glutamate binds to two types of receptors: metabotropic (mGluRs) and ionotropic glutamate receptors (iGluRs). iGluRs can be divided into N-methyl-D-aspartate (NMDA) and non-NMDA receptors (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid [AMPA] and kainate receptors) according to their response to agonist molecules NMDA and AMPA (Mosbacher et al., 1994). Analysis of the *Drosophila* genome annotated 14 iGluRs genes, which show sequence similarities with vertebrate AMPA, kainite, and NMDA receptors (Littleton and Ganetzky, 2000). However, the kainite receptor DKaiR1D and the AMPA receptor DGluR1A have different agonist/antagonist selectivity from the vertebrate's pharmacology-based classification (Li et al., 2016). Furthermore, invertebrates like *Drosophila melanogaster* possess a third type of iGluR, the so-called glutamate-gated chloride channel GluClα, which

is inhibitory (Cully et al., 1996; Liu and Wilson, 2013). Glutamate can also act on metabotropic glutamate receptors, which signal via slower G-protein-coupled pathways. In mammals, eight mGluRs have been described (Conn and Pin, 1997). In contrast, the *Drosophila* genome encodes only one functional mGluR (DmGluRA), which is expressed at the glutamatergic NMJ localized in the presynaptic boutons (Bogdanik et al., 2004). Regarding the broad range of glutamate receptors in *Drosophila*, glutamate can act as a fast, slow, excitatory, or inhibitory transmitter (Li et al., 2016; Liu and Wilson, 2013; Mauss et al., 2015).

This gives rise to interesting speculations about the respective role of glutamate for each of the cell types investigated. In the case of the LPis, glutamate binds to the inhibitory glutamate receptor GluClα on the dendrites of large-field tangential cells, inhibiting them during null direction motion and, thus, enhancing their flow-field selectivity (Mauss et al., 2015). In the case of L1, the glutamatergic output signal seems to be key for the sign inversion of L1's OFF response in the ON pathway. This is because all *Drosophila* photoreceptors (R1-R8) depolarize upon illumination and release histamine onto lamina neurons, which results in the opening of chloride channels (Hardie, 1989; Hardie and Raghu, 2001). Therefore, lamina monopolar cells transiently hyperpolarize upon illumination onset and respond with a rebound excitation at illumination offset (Laughlin et al., 1987). L1 and L2 neurons respond in an identical way (Joesch et al., 2010). L1 possess an OFF receptive field center (Figure 3D) and therefore depolarizes to OFF stimuli, in contrast to its described downstream synaptic partners, which depolarize to ON stimuli (Arenz et al., 2017; Behnia et al., 2014; Strother et al., 2017; Yang et al., 2016). Hence, an inversion of the sign must occur at the synapse of L1 and its downstream partners. Since L1 is glutamatergic and GluClα is the only inhibitory receptor described in *Drosophila*, the glutamatergic signal is likely to be responsible for this sign inversion. Whether the downstream partners of L1 indeed express GluClα, however, is beyond the scope of this study and awaits further investigation. The hypothesis outlined above suggests that the mechanism by which a common photoreceptor input signal is split into an ON and an OFF pathway in invertebrates is different from the one in the mammalian retina where glutamatergic photoreceptors hyperpolarize in response to light. This signal is directly transmitted, i.e., without sign inversion, by ionotropic glutamate receptors expressed on the dendrites of OFF bipolar cells (Euler et al., 2014) and sign inverted by metabotropic glutamate receptors expressed on the dendrites of ON bipolar cells (Masu et al., 1995). In case of Mi9, the functional interpretation of an inhibitory glutamatergic signal is less intuitive. Mi9 directly contacts the dendrites of T4 cells, the first direction-selective neurons in the ON pathway (Takemura et al., 2017). Given the OFF response of Mi9 cells (Figure 3D), T4 cells are expected to be inhibited in darkness via the Mi9-T4 synapse. A moving ON edge would inhibit Mi9 followed by a closure of chloride channels and, thus, an increased input resistance in postsynaptic T4 cells, resulting in an amplification of a subsequently delivered excitatory input signal. Computer simulations have shown that such a two-fold signal inversion can indeed form the biophysical basis of preferred direction enhancement underlying direction selectivity in T4 cells (Borst, 2018).

Taken together our results could demonstrate the functionality of the fast glutamate reporter iGluSnFR in glutamatergic neurons of the fruit fly *Drosophila melanogaster*. It allowed for a more faithful description of important elements of the motion vision pathway, in particular with respect to their temporal response properties.

Limitations of the Study

Since iGluSnFR is anchored to the outer side of the plasma membrane, it senses extracellular glutamate that is present in the synaptic cleft. In addition, the iGluSnFR signal is affected by spillover and diffusion to iGluSnFR molecules outside the cleft. Thus, the iGluSnFR signal should present an upper limit to the "real" time course, i.e., the one of glutamate in the synaptic cleft as seen by the postsynaptic receptors. For the same reason, one might record an iGluSnFR signal even if the indicator is expressed on a neuron that is not glutamatergic or does not receive glutamatergic input, but ramifies within the same volume where glutamate is being released from other cells.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods and two figures and can be found with this article online at https://doi.org/10.1016/j.isci.2018.08.019.

ACKNOWLEDGMENTS

We thank Aljoscha Leonhardt for careful proofreading of the manuscript. We would also like to acknowledge Hermann Aberle for sharing the VGlut antibody with us and Julia Kuhl for designing the graphical abstract. We thank Wolfgang Essbauer and Michael Sauter for fly husbandry. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 870) and the Max Planck Society. F.G.R., M.S.D., S.F., and A.B. are members of the Graduate School of Systemic Neurosciences (GSN) Munich.

AUTHOR CONTRIBUTIONS

F.G.R., S.F., and A.B. conceived the study and designed the experiments. F.G.R. conducted and analyzed the imaging experiments for Mi9 and L1. S.F. performed and analyzed all stainings. J.H. performed and analyzed the LPi experiments. M.S.D. performed data analysis and model fitting of the receptive fields. F.G.R. wrote the manuscript with the help of all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 5, 2018
Revised: August 13, 2018
Accepted: August 23, 2018
Published: September 28, 2018

REFERENCES

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.

Barlow, H.B., and Levick, W.R. (1965). The mechanism of directionally selective units in rabbit's retina. J. Physiol. *178*, 477–504.

Bausenwein, B., Dittrich, A.P.M., and Fischbach, K.F. (1992). The optic lobe of *Drosophila melanogaster* - II. Sorting of retinotopic pathways in the medulla. Cell Tissue Res. *267*, 17–28.

Bausenwein, B., and Fischbach, K.F. (1992). Activity labeling patterns in the medulla of *Drosophila melanogaster* caused by motion stimuli. Cell Tissue Res. *270*, 25–35.

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.

Bogdanik, L., Mohrmann, R., Ramaekers, A., Bockaert, J., Grau, Y., Broadie, K., and Parmentier, M.L. (2004). The *Drosophila* metabotropic glutamate receptor DmGluRA regulates activity-dependent synaptic facilitation and fine synaptic morphology. J. Neurosci. *24*, 9105–9116.

Borst, A. (2018). A biophysical mechanism for preferred direction enhancement in fly motion vision. PLoS Comput. Biol. *14* (6), e1006240.

Borst, A. (2014). In search of the holy grail of fly motion vision. Eur. J. Neurosci. *40*, 3285–3293.

Borst, A., and Abarbanel, H.D. (2007). Relating a calcium indicator signal to the unperturbed calcium concentration time-course. Theor. Biol. Med. Model. *4*, 7.

Cao, G., Platisa, J., Pieribone, V.A., Raccuglia, D., Kunst, M., and Nitabach, M.N. (2013). Genetically targeted optical electrophysiology in intact neural circuits. Cell *154*, 904–913.

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.

Chichilnisky, E.J. (2001). A simple white noise analysis of neuronal light responses. Network *12*, 199–213.

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.

Conn, P.J., and Pin, J.-P. (1997). Pharmacology and function of metabotropic glutamate receptors. Annu. Rev. Pharmacol. Toxicol. *37*, 205–237.

Cruse, H. (1996). Neural Networks as Cybernetics Systems.

Cully, D.F., Paress, P.S., Liu, K.K., Schaeffer, J.M., and Arena, J.P. (1996). Identification of a *Drosophila melanogaster* glutamate-gated chloride channel sensitive to the antiparasitic agent avermectin. J. Biol. Chem. *271*, 20187–20191.

Daniels, R.W. (2004). Increased expression of the *Drosophila* vesicular glutamate transporter leads to excess glutamate release and a compensatory decrease in quantal content. J. Neurosci. *24*, 10466–10474.

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/10.1101/385476.

Dayan, P., and Abbott, L.F. (2013). Theoretical neuroscience: computational and mathematical modeling of neural systems. J. Chem. Inf. Model. *53*, 1689–1699.

Di Maio, V. (2008). Regulation of information passing by synaptic transmission: a short review. Brain Res. *1225*, 26–38.

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.

Euler, T., Haverkamp, S., Schubert, T., and Baden, T. (2014). Retinal bipolar cells: elementary building blocks of vision. Nat. Rev. Neurosci. *15*, 507–519.

French, A.S. (1976). Practical nonlinear system analysis by Wiener kernel estimation in the frequency domain. Biol. Cybern. *24*, 111–119.

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.

Haag, J., Arenz, A., Serbe, E., Gabbiani, F., and Borst, A. (2016). Complementary mechanisms create direction selectivity in the fly. Elife *5*, https://doi.org/10.7554/eLife.17421e.

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*, 1–15.

Hardie, R.C. (1989). A histamine-activated chloride channel involved in neurotransmission at a photoreceptor synapse. Nature *339*, 704–706.

Hardie, R.C., and Raghu, P. (2001). Visual transduction in *Drosophila*. Nature *413*, 186–193.

Hausen, K., Wolburg-Buchholz, K., and Ribi, W.A. (1980). The synaptic organization of visual interneurons in the lobula complex of flies - a light and electron microscopical study using silver-intensified cobalt-impregnations. Cell Tissue Res. *208*, 371–387.

Hopp, E., Borst, A., and Haag, J. (2014). Subcellular mapping of dendritic activity in optic flow processing neurons. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *200*, 359–370.

Jin, L., Han, Z., Platisa, J., Wooltorton, J.R.A., Cohen, L.B., and Pieribone, V.A. (2012). Single action potentials and subthreshold electrical events imaged in neurons with a fluorescent protein voltage probe. Neuron *75*, 779–785.

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.

Joesch, M., Weber, F., Eichner, H., and Borst, A. (2013). Functional specialization of parallel motion detection circuits in the fly. J. Neurosci. *33*, 902–905.

Juusola, M., Dau, A., Zheng, L., and Rien, D. (2016). Electrophysiological method for recording intracellular voltage responses of *Drosophila* photoreceptors and interneurons to light stimuli in vivo. J. Vis. Exp. *112*, 1–16.

Kolodziejczyk, A., Sun, X., Meinertzhagen, I.A., and Nässel, D.R. (2008). Glutamate, GABA and acetylcholine signaling components in the lamina of the *Drosophila* visual system. PLoS One *3*, e2110.

Laughlin, S.B., Howard, J., and Blakeslee, B. (1987). Synaptic limitations to contrast coding in the retina of the blowfly *Calliphora*. Proc. R. Soc. Lond. B Biol. Sci. *231*, 437–467.

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.

Li, Y., Dharkar, P., Han, T.H., Serpe, M., Lee, C.H., and Mayer, M.L. (2016). Novel functional properties of *Drosophila* CNS glutamate receptors. Neuron *92*, 1036–1048.

Littleton, J.T., and Ganetzky, B. (2000). Ion channels and synaptic organization: analysis of the *Drosophila* genome. Neuron *26*, 35–43.

Liu, W.W., and Wilson, R.I. (2013). Glutamate is an inhibitory neurotransmitter in the *Drosophila* olfactory system. Proc. Natl. Acad. Sci. USA *110*, 10294–10299.

Mahr, A., and Aberle, H. (2006). The expression pattern of the *Drosophila* vesicular glutamate transporter: a marker protein for motoneurons and glutamatergic centers in the brain. Gene Expr. Patterns *6*, 299–309.

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.

Marvin, J.S., Borghuis, B.G., Tian, L., Cichon, J., Harnett, M.T., Akerboom, J., Gordus, A., Renninger, S.L., Chen, T.W., Bargmann, C.I., et al. (2013). An optimized fluorescent probe for visualizing glutamate neurotransmission. Nat. Methods *10*, 162–170.

Masu, M., Iwakabe, H., Tagawa, Y., Miyoshi, T., Yamashita, M., Fukuda, Y., Sasaki, H., Hiroi, K., Nakamura, Y., Shigemoto, R., et al. (1995). Specific deficit of the ON response in visual transmission by targeted disruption of the mGluR6 gene. Cell *80*, 757–765.

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.

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.

Meldrum, B.S. (2000). Glutamate as a neurotransmitter in the brain: review of physiology and pathology. J. Nutr. *130*, 1007S–1015S.

Mosbacher, J., Schoepfer, R., Monyer, H., Burnashev, N., Seeburg, P., and Ruppersberg, J. (1994). A molecular determinant for submillisecond desensitization in glutamate receptors. Science *266*, 1059–1062.

Raghu, S.V., and Borst, A. (2011). Candidate glutamatergic neurons in the visual system of *Drosophila*. PLoS One. https://doi.org/10.1371/journal.pone.0019472.

Ringach, D., and Shapley, R. (2004). Reverse correlation in neurophysiology. Cognit. Sci. *28*, 147–166.

Ringach, D.L. (2004). Mapping receptive fields in primary visual cortex. J. Physiol. *558*, 717–728.

Rister, J., Pauls, D., Schnell, B., Ting, C.Y., Lee, C.H., Sinakevitch, I., Morante, J., Strausfeld, N.J., Ito, K., and Heisenberg, M. (2007). Dissection of the peripheral motion channel in the visual system of *Drosophila melanogaster*. Neuron *56*, 155–170.

Schnell, B., Joesch, M., Forstner, F., Raghu, S.V., Otsuna, H., Ito, K., Borst, A., and Reiff, D.F. (2010). Processing of horizontal optic flow in three visual interneurons of the *Drosophila* brain. J. Neurophysiol. *103*, 1646–1657.

Scott, E.K., Raabe, T., and Luo, L. (2002). Structure of the vertical and horizontal system neurons of the lobula plate in *Drosophila*. J. Comp. Neurol. *454*, 470–481.

Shapley, R., and Lennie, P. (1985). Spatial frequency analysis in the visual system. Annu. Rev. Neurosci. *8*, 547–583.

Shinomiya, K., Karuppudurai, T., Lin, T.Y., Lu, Z., Lee, C.H., and Meinertzhagen, I.A. (2014). Candidate neural substrates for OFF edge motion detection in *Drosophila*. Curr. Biol. *24*, 1062–1070.

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.

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*, 1471–2954.

St-Pierre, F., Marshall, J.D., Yang, Y., Gong, Y., Schnitzer, M.J., and Lin, M.Z. (2014). High-fidelity optical reporting of neuronal electrical activity with an ultrafast fluorescent voltage sensor. Nat. Neurosci. *17*, 884–889.

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.

Takemura, S.Y., Nern, A., Chklovskii, D.B., Scheffer, L.K., Rubin, G.M., and Meinertzhagen, I.A. (2017). The comprehensive connectome of a neural substrate for 'ON' motion detection in *Drosophila*. Elife *6*, 1–16.

Takemura, S.Y., Karuppudurai, T., Ting, C.Y., Lu, Z., Lee, C.H., and Meinertzhagen, I.A. (2011). Cholinergic circuits integrate neighboring visual signals in a *Drosophila* motion detection pathway. Curr. Biol. *21*, 2077–2084.

Traynelis, S.F., Wollmuth, L.P., McBain, C.J., Menniti, F.S., Vance, K.M., Ogden, K.K., Hansen, K.B., Yuan, H., Myers, S.J., and Dingledine, R. (2010). Glutamate receptor ion channels: structure, regulation, and function. Pharmacol. Rev. *62*, 405–496.

Tsutsui, H., Jinno, Y., Tomita, A., Niino, Y., Yamada, Y., Mikoshiba, K., Miyawaki, A., and Okamura, Y. (2013). Improved detection of electrical activity with a voltage probe based on a voltage-sensing phosphatase. J. Physiol. *591*, 4427–4437.

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.

Von Hassenstein, B., and Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen-und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers Chlorophanus. Sect. B J. Chem. Sci. *11*, 513–524.

Wasserman, S.M., Aptekar, J.W., Lu, P., Nguyen, J., Wang, A.L., Keles, M.F., Grygoruk, A., Krantz, D.E., Larsen, C., and Frye, M.A. (2015). Olfactory neuromodulation of motion vision circuitry in *Drosophila*. Curr. Biol. *25*, 467–472.

Yang, H.H., St-Pierre, F., Sun, X., Ding, X., Lin, M.Z., and Clandinin, T.R. (2016). Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. Cell *166*, 245–257.

Zucker, R.S. (1993). Calcium and transmitter release. J. Physiol. Paris *87*, 25–36.

ISCI, Volume 7

Supplemental Information

Glutamate Signaling in the Fly Visual System

Florian G. Richter, Sandra Fendl, Jürgen Haag, Michael S. Drews, and Alexander Borst

Supplemental Information

Supplemental Figures

A

L1	GCaMP6f	iGluSnFR
τ_{LP}	350 ms	70 ms
τ_{HP}	1178 ms	407 ms
$FWHM_{cen}$	8.74	10.56
$FWHM_{sur}$	56.55	42.55
A_{rel}	19.33	20.51

B

Mi9	GCaMP6f	iGluSnFR
τ_{LP}	607 ms	75 ms
$FWHM_{cen}$	6.69	7.18
$FWHM_{sur}$	19.37	25.86
A_{rel}	2.81	3.56

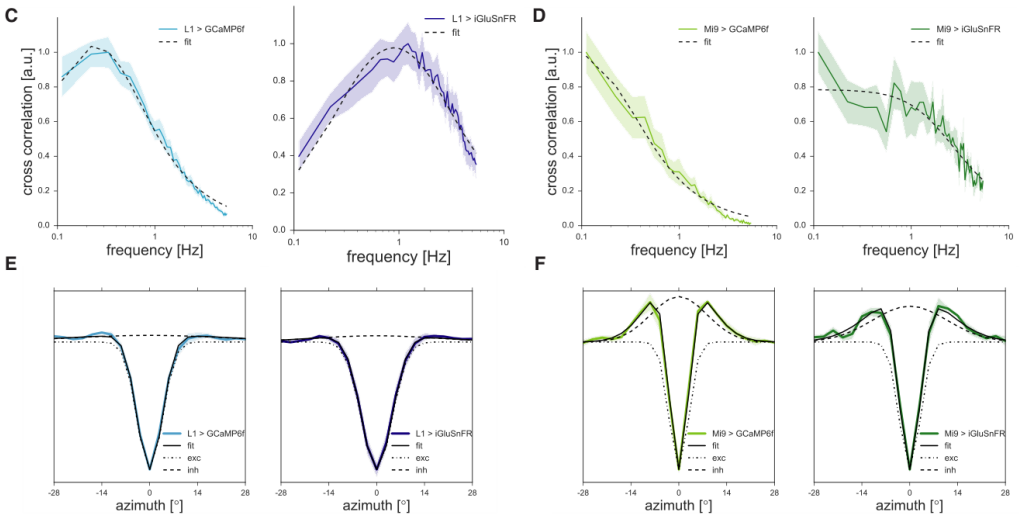


Figure S1. Model fits to L1 and Mi9 data, related to Fig 3
(A) Parameters to quantitatively describe the receptive field characteristics of L1 recorded either with GCaMP6f (left column) or iGluSnFR (right column). First two parameters describe temporal components of the receptive field, last three parameters describe those of the spatial component.
(B) Same as (A) only for Mi9. Description of highpass characteristics is missing, since Mi9 is best described by a pure low-pass.
(C) Impulse responses from Figure 3 D-E plotted in frequency space. Black dashed lines mark the fit of a 1st order band-pass filter (for time constants see table (A).
(D) Same as (C) only for Mi9. Black dashed lines mark the fit of a 1st order low-pass filter.
(E)+(F) Spatial receptive fields from Figure 3 D-E. Data are fitted with a Mexican hat function that captures both, the excitatory center as well as the inhibitory surround of these receptive fields. cen = center, sur = surround, LP = low-pass, HP = high-pass, A = amplitude, τ = time constant, FHWM = full width at half maximum.

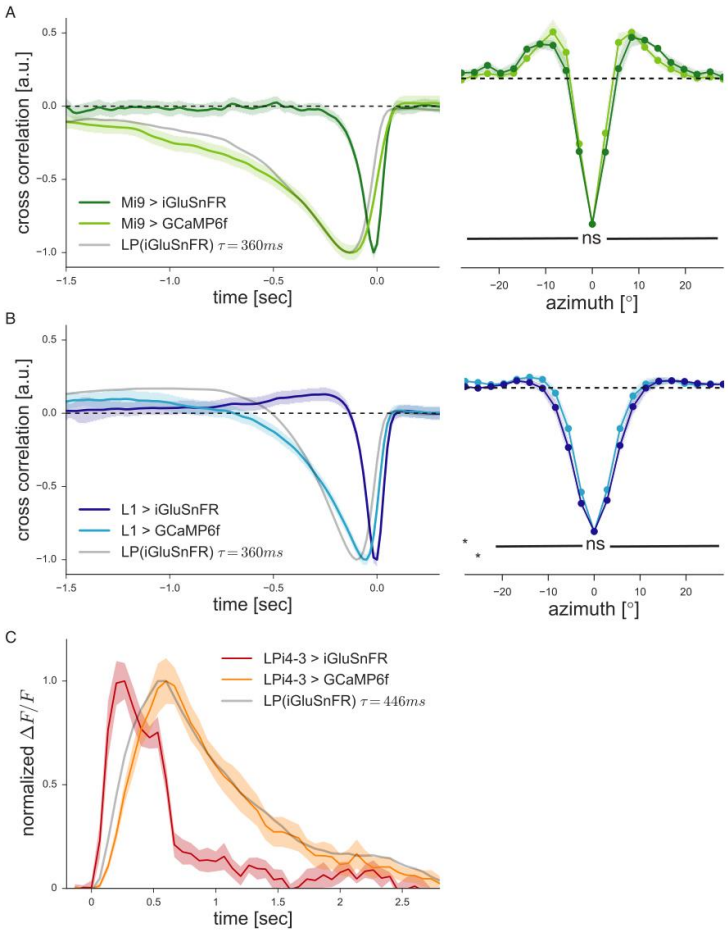


Figure S2. GCaMP data resembles low-pass filtered iGluSnFR data, related to Fig 3 and 4
(A) Low-pass filtering of the Mi9 impulse response measured with iGluSnFR with a time constant of 360 ms (grey) shows the best fit with the impulse response measured with GCaMP6f (left panel). Spatial receptive fields (right panel) are not significantly different from each other, when measured with the two different sensors.
(B) Same as (A) for L1
(C) Low-pass filtering of the LPi4-3 > iGluSnFR response to local flicker with a time constant of 446 ms (grey) shows the best fit to response measured with GCaMP6f (orange).

Transparent Methods

Flies/preparation

Flies were raised and kept on standard cornmeal-agar medium on a 12 hour light/12 hour dark cycle at 25°C and 60% humidity. For imaging experiments, the genetically-encoded calcium indicators GCaMP6f or the genetically encoded glutamate sensor iGluSnFR (Chen et al., 2013; Marvin et al., 2013) were expressed using the Gal4-UAS system in cell-type specific Gal4 fly lines, resulting in the following genotypes:

Genotypes:

L1>GC6f:	w+; R48A08-AD/UAS-GCaMP6f; R66A01-DBD/UAS-GCaMP6f
L1>iGluSnFR:	w+; R48A08-AD/+; R66A01-DBD/UAS-iGluSnFR (BL59611, AV184)
Mi9>GC6f:	w+; R48A07-AD/UAS-GCaMP6f; VT046779-DBD/UAS-GCaMP6f
Mi9>iGluSnFR:	w+; R48A07-AD/+; VT046779-DBD/UAS-iGluSnFR (BL59611, AV184)
LPi>GC6f:	w+; +/UAS-GCaMP6f; R38G02-Gal4/UAS-GCaMP6f
LPi>iGluSnFR:	w+; +; R38G02-Gal4/UAS-iGluSnFR (BL59611, AV184)

For immunohistochemical stainings in Figure 2:

L1>myr::GFP:	w-; R48A08-AD/UAS-myr::GFP; R66A01-DBD/+
Mi9>myr::GFP:	w-; R48A07-AD/ UAS-myr::GFP; VT046779-DBD/+
LPi4-3>myr::GFP:	w-; UAS-myr::GFP/+; R38G02-Gal4/+

The transgenic fly lines driving split-Gal4 expression in the lamina neuron L1 were generated and described in (Tuthill et al., 2013). Mi9 in (Strother et al., 2017) and the one of LPi's in (Mauss et al., 2015). For calcium and glutamate imaging experiments, flies were prepared as previously described (Maisak et al., 2013; Strother et al., 2017). Briefly, flies were anaesthetized on ice, fixed with their backs, legs and wings to a Plexiglas holder with the back of the head exposed to a recording chamber filled with fly external solution. The cuticle at the back of the head on one side was cut away with a fine hypodermic needle and removed together with muscles and air sacks covering the underlying optic lobe.

Data acquisition and analysis:

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

2-photon imaging:

Imaging was performed on custom-built 2-photon microscopes as previously described (Maisak et al., 2013) and controlled with the ScanImage software in Matlab (Pologruto et al., 2003). Acquisition rates were between 15 (for LPi experiments) and 23.67 Hz (for L1 and Mi9 experiments), image resolution between 64x64 and 128x32 pixels (for L1 and Mi9 experiments). Before starting the acquisition, we verified that the receptive fields of the cells were located on the stimulus arena by showing a search stimulus consisting of moving gratings.

Calcium imaging was performed as previously described in (Arenz et al., 2017). In brief: Images were automatically registered using horizontal and vertical translations to correct for the movement of the brain. Fluorescence changes (ΔF/F values) were then calculated using a standard baseline algorithm (Jia et al., 2011). Regions of interest (ROIs) were drawn on the average raw image by hand in the medulla layer M1 for L1 and in layer M10 for Mi9. For LPi neurons, ROIs were routinely chosen in the lobula plate, encompassing small regions with single to few axon terminals. Averaging the fluorescence change over this ROI in space resulted in a ΔF/F time course. Glutamate imaging was performed with the same settings as the calcium imaging experiments.

Visual stimulation for L1 and Mi9 experiments

The spatiotemporal response properties of the L1 and Mi9 columnar input elements were determined on a custom-built projector-based arena, as previously described in (Arenz et al., 2017). Stimuli were projected with 2 commercial micro-projectors (TI DLP Lightcrafter 3000) onto the back of an opaque cylindrical screen covering 180 ° in azimuth and 105 ° in elevation of the fly's visual field. The projectors refresh rate is 180 Hz (at 8 bit color depth). For all stimuli used here, we set the medium brightness to a 8-bit grayscale value of 50, which corresponds to a medium luminance of 55 ± 11 cd/m^2. Stimuli were rendered using a custom written software in Python 2.7.

Visual stimulation for LPi4-3 experiments with telescope

This technique has been previously described in (Haag et al., 2016). In brief: Antidromic illumination of the fly's head visualizes the hexagonal structure of the optical axes of the ommatidia (Franceschini, 1975; Schuling et al., 1989). Visual stimuli are generated on the AMOLED display (800x600 pixels, pixel size 15x15 mm, maximal luminance > 1500 cd/m^2; lambda = 530 nm; refresh rate 85 Hz) (SVGA050SG, Olightek). This allows to precisely position the stimuli onto single lamina cartriges. In order to prevent stimulus light from entering the photomultiplier of the two-photon micro-scope, light generated by the AMOLED display was filtered with a long-pass filter (514 LP, T: 529.4– 900 nm, AHF). The AMOLED display was controlled with MATLAB and the psychophysics toolbox (V3.0.11;(Brainard, 1997)).

White noise reverse-correlation

The analysis of spatial receptive fields was previously described in (Arenz et al., 2017). For the input elements, spatiotemporal receptive fields were calculated following standard reverse-correlation methods (Dayan and Abbott, 2013; French, 1976). First, the mean value was subtracted from the raw signals of single ROIs by using a low-pass filtered version of the signal (Gaussian filter with 120 seconds standard deviation) as a baseline for a ΔF/F-like representation of the signal.
The stimulus-response reverse correlation function was calculated as:

$$K(x,\tau) = \int_0^T dt\, S(x, t-\tau) \cdot R(t)$$

with S for the stimulus and R for the response of the neuron. The resulting spatiotemporal fields were normalized in z-score. Only receptive fields with peak amplitudes above 10 standard deviations from the mean were taken for further analysis (for Mi9-GCaMP6f the threshold

was lowered to 7). Cross-sections through the receptive fields along the space axis were fit with a Gaussian function to determine the position of the peak (Suppl. Fig. 1 E-F).

Gaussian noise stimulus

The same stimulus was used in (Arenz et al., 2017). In brief: The stimulus consisted of 64 vertical bars covering an angle of 180° in total. The intensity of each bar fluctuated randomly around a mean intensity of 50 on the 8-bit grayscale of the display. The intensities were drawn from a Gaussian distribution with a standard deviation of 25% contrast. In time, the stimulus was low-pass filtered with a Gaussian window with approximately 22ms standard deviation, which restricted the frequency content of the stimulus to frequencies below 10Hz. For Mi9-GCaMP6f imaging, similarly, the time window was 45ms long, covering frequencies until up to 5Hz.

Spatial receptive field

The analysis of spatial receptive fields was previously described in (Arenz et al., 2017). In brief: One-dimensional spatial receptive fields are cross-sections through the peak of the spatiotemporal receptive fields along the space axis and are averaged over the 12 samples (200ms) around the peak. For both L1 and Mi9 we found a small-field, antagonistic center-surround organization of the spatial receptive field using the vertical white noise stimulus.
The black dashed lines in Suppl. Fig 1 represents a Mexican hat function (Difference of Gaussian). Mathematically such a function can be described as follows:

$$RF_{1D}(\varphi) = e^{-\frac{1}{2}\frac{\varphi^2}{\sigma_{cen}{}^2}} - A_{rel} \cdot e^{-\frac{1}{2}\frac{\varphi^2}{\sigma_{sur}{}^2}}$$

with φ as azimuth, σ_{cen} and σ_{sur} as the standard deviations of center and surround, respectively, and $A_{rel} = A_{sur}/A_{cen}$ the relative strength of the surround in relation to the amplitude of the center Gaussian (which is normalized to 1).

Temporal receptive field

The analysis of temporal receptive fields was previously described in (Arenz et al., 2017). In brief: The time-reversed impulse responses shown in Figure 3 are cross-sections through the center of the spatiotemporal receptive fields along the time axis and are averaged over the three center pixels. For the determination of the time constants (tau), we sought to describe the response characteristic of each cell with a simplified model that catches the main properties. For that, we fitted simple 1st-order filters (e.g. 1st order low-pass for Mi9; 1st order bandpass for L1) to the impulse responses of all cells.
The model fit in Suppl. Fig 2 (grey lines) was performed by low-pass filtering the measured iGluSnFR response of each neuron type (L1, Mi9, LPi) with a 1st order low-pass filter and optimizing the time-constant such that the difference between the low-pass filtered signal and the measured calcium response of the neurons was minimal. The fitting procedure was implemented using standard least square algorithms (SciPy 0.19).

Immunohistochemistry

Fly brains were dissected in ice-cold 0.3% PBST and fixed in 4% PFA in 0.3% PBST for 25 min at room temperature. Subsequently, brains were washed 4-5 times in 0.3% PBST and blocked in 10% normal goat serum (NGS) in 0.3% PBST for 1 hour at room temperature. Primary antibodies used were mouse anti-bruchpilot brp (nc82, Developmental Studies

Hybridoma Bank, 1:20) and rabbit anti-VGlut (courtesy of H. Aberle, 1:500). Secondary antibodies used were: goat anti-mouse ATTO 647N (Rockland, 1:300) and goat anti-rabbit Alexa Fluor 568 (Life Technologies, 1:300). Myr::GFP-labeled cells were imaged natively without antibody staining. 5% NGS was added to all antibody solutions and both primary and secondary antibodies were incubated for at least 48 hours at 4°C.
Brains were mounted in Vectashield Antifade Mounting Medium (Vector Laboratories) and imaged on a Leica TCS SP8 confocal microscope.

Supplemental References

Brainard, D.H. (1997). The psychophysics toolbox. Spat. Vis. 10, 433–436.

Franceschini, N. (1975). Sampling of the visual environment by the compound eye of the fly: Fundamentals and applications. Photoreceptor Opt. 98–125.

Jia, H., Rochefort, N.L., Chen, X., Konnerth, A. (2011). In vivo two-photon imaging of sensory-evoked dendritic calcium signals in cortical neurons. Nat. Protoc. 6, 28–35.

Pologruto, T.A., Sabatini, B.L., Svoboda, K. (2003). ScanImage: Flexible software for operating laser scanning microscopes. Biomed. Eng. Online 2.

Schuling, F.H., Mastebroek, H.A.K., Bult, R., Lenting, B.P.M. (1989). Properties of elementary movement detectors in the fly *Calliphora erythrocephala*. J. Comp. Physiol. A 165, 179–192.

2.4 DYNAMIC SIGNAL COMPRESSION FOR ROBUST MOTION VISION IN FLIES

SUMMARY This study is dedicated to the investigation of non-linear contrast tuning properties in the fly motion vision circuitry and shows that divisive contrast gain control in medulla neurons is crucial for robust motion processing of naturalistic input stimuli.

In previous studies, we had investigated motion processing pathways in the fly by use of artificial stimuli with constant contrast. However, it is known that standard models of motion detection are extremely vulnerable to the contrast fluctuations inherent to natural input stimuli. In this study, we first established that flies robustly perform velocity estimation from moving natural scenes regardless of the particular visual statistics of the stimulus. We then used behavioral experiments to show that flies take advantage of a spatial contrast gain control mechanism to adjust the contrast sensitivity of motion processing pathways to the prevailing contrast of the visual scene. Extensive calcium imaging experiments across all columnar cell types in the fly motion vision circuitry revealed that contrast gain control emerges in a subset of medulla neurons. Specifically, transient band-pass filter units seem to implement a kind of contrast gain control that resembles the mechanism of divisive normalization, which has been previously described in the vertebrate visual system. Using blocking experiments, we could show that this mechanism partially originates from feedback connections of transient medulla neurons onto themselves. Finally, we built a convolutional network model of the fly motion vision circuitry to show that a motion vision system equipped with this kind of dynamic gain control is superior to other types of non-linear signal preprocessing at estimating scene velocity from naturalistic input stimuli and achieves fly-like performance at this task.

This manuscript is currently submitted to a peer-reviewed journal.

AUTHORS **Michael S. Drews**, Aljoscha Leonhardt (co-first author), Nadezhda Pirogova, Florian G. Richter, Anna Schuetzenberger, Lukas Braun, Etienne Serbe, & Alexander Borst

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 modelling, and wrote the manuscript. All authors participated in editing the manuscript.

Dynamic signal compression for robust motion vision in flies

Michael S. Drews[1,2,4], Aljoscha Leonhardt[1,4], Nadezhda Pirogova[1,2], Florian G. Richter[1,2], Anna Schuetzenberger[1,2], Lukas Braun[3], Etienne Serbe[1], and Alexander Borst[1]

[1]Department Circuits-Computation-Models, Max-Planck-Institute of Neurobiology, Martinsried, Germany
[2]Graduate School of Systemic Neurosciences, Ludwig-Maximilians-Universität, Munich, Germany
[3]Bernstein Center for Computational Neuroscience, Berlin, Germany
[4]These authors contributed equally: Michael S. Drews, Aljoscha Leonhardt

Biological and artificial sensory systems need to reliably extract information from highly variable natural signals[1, 2]. Flies, for instance, use optic flow to guide their course[3] and are remarkably adept at estimating image velocity regardless of visual statistics in any given environment[4, 5]. It remains unclear, however, how this robustness is achieved. Current models of fly motion detection explain responses to artificial stimuli in great detail but are vulnerable to the extreme contrast fluctuations pervasive in realistic images. Here, we demonstrate that the Drosophila visual system reduces input variability by continuously adjusting its sensitivity to current contrast conditions. We comprehensively map functional properties of neurons in the motion detection circuit and find that local responses are nonlinearly compressed by a signal that estimates surround contrast. The compressive signal is fast, integrates spatially, and derives from neural feedback. This mechanism resembles divisive normalization as commonly found in vertebrate visual processing[6, 7], emphasizing that evolutionarily distant neural systems often converge on similar algorithmic solutions. We train fly-like convolutional neural networks[8] on estimating the velocity of natural stimuli and show that introducing dynamic signal compression closes the performance gap between model and organism. This may provide a building block for the efficient implementation of low-power machine vision systems. Overall, our work represents a mechanistic end-to-end account of how neural systems attain the robustness required to carry out behaviorally relevant tasks in challenging real-world environments.

Correspondence: ***leonhardt@neuro.mpg.de, drews@neuro.mpg.de***

To rigorously assess the robustness of Drosophila motion processing we took advantage of a widespread visual reflex, the optomotor response. Flies react to whole-field retinal motion by turning in the same direction as their surround which allows them to maintain a straight path under perturbations[3]. For this reflex to work effectively, flies need to respond reliably and independently of the particular visual statistics of their environment. We measured walking responses to a diverse set of moving naturalistic panoramas on a treadmill setup (Fig. 1a). Fly turning was highly consistent across images and velocity tuning curves showed virtually no variation over different scenes, matching previous findings[4] (Fig. 1b, Fig. S1a–c). To quantify reliability at the neural level, we recorded the membrane potential of motion-sensitive lobula plate tangential cells (LPTC) which detect optic flow fields corresponding to yaw rotation and control turning[3] (Fig. 1c). LPTCs pool the output of a retinotopic map of locally motion-sensitive cell types T4 and T5 that

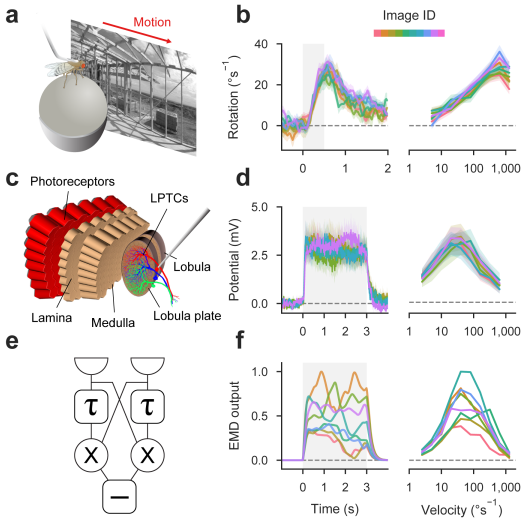


Figure 1. Flies respond more robustly to natural scene variability than predicted by correlation-based motion detectors. a, Illustration of behavioral setup. Tethered wild-type Drosophila were stimulated with translating natural images. **b,** Left, turning responses for images moving at 80°s^{-1} (N=16 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 1s after motion onset). **c,** Illustration of fly visual system. Photoreceptor signals are processed in five retinotopically arranged neuropils. Wide-field lobula plate tangential cells (LPTC) 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 3s after motion onset). **e,** Schematic of an individual correlation-based elementary motion detector (EMD; τ denotes delay line, X 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, mean responses were averaged across many different starting phases. Shaded areas around curves in b,d indicate bootstrapped 68% confidence intervals.

are sensitive to ON and OFF motion[9], respectively, and necessary for the optomotor response[4, 10]. Voltage was tuned to scene velocity but again exhibited little image-dependent variation and additionally proved highly stable across time, consistent with earlier work in hoverflies[5] (Fig. 1d).

Fly motion responses are well explained by correlation-based detector models that rely on multiplication of spatially adjacent, asymmetrically filtered luminance signals[3, 11] (Fig. 1e). These elementary motion detectors (EMD) account for subtle features of optomotor responses, LPTCs, and T4/T5 cells

such as pattern dependency of velocity tuning, reverse-phi sensitivity, or velocity gain adaptation[3, 12, 13]. We tested the robustness of EMDs on the same set of stimuli as in behavior and electrophysiology. As anticipated from similar studies[4, 14, 15, 16], responses were remarkably unreliable across time and images (Fig. 1f). EMD output strongly depends on local contrast as defined by the average difference between light and dark[11]. However, contrast within natural images varies substantially[2]. Correlation-based models invariably confound velocity with contrast and consequently, output from individual EMDs is sparse and fluctuates heavily under naturalistic conditions (Fig. S1d–f). This stands in stark contrast to the observed robustness of motion responses and leads to the central question: How does the fly visual system compensate for natural contrast variability?

Various general mechanisms for adaptation to naturalistic signals have been described in the fly visual system including gain control in photoreceptors or LPTCs[17, 18, 19], redundancy reduction through lateral inhibition[20], subtractive enhancement of flow field selectivity[3], and tailoring of processing to fundamental natural scene statistics[4, 14]. However, none effectively addresses the problem of contrast fluctuations.

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 segregates the visual field into a background pattern containing random luminance fluctuations at a controlled contrast level and a random foreground pattern whose motion triggers turning (Fig. 2a). We confirmed that background produced no net activity in EMDs (Fig. 2b). At zero background contrast, foreground motion induced a reliable optomotor response (Fig. 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 surround contrast (Fig. 2d,e).

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 (Fig. 2f). Whenever background contrast was high, syndirectional rotation in response to motion was transiently suppressed. Evaluating oscillation amplitude thus allowed a read-out of the level of contrast-induced gain adjustment. We determined the spatial scale by varying spacing between fore- and background. Modulation fell with distance and dropped to baseline at approximately 40° of separation, so contrast estimation was non-local but spatially limited (Fig. 2g, Fig. S2a,b). When we varied oscillation frequency, suppression followed contrast changes up to fast time scales beyond 3 Hz (Fig. 2h, Fig. S2c–f). However, modulation decreased at lower frequencies than for equivalent foreground oscillations which is indicative of temporal integration. Silencing T4 and T5 cells abolished all contrast-guided oscillatory turning (Fig. S2g–i), suggesting that contrast adaptation is not mediated by a system parallel to motion detection[10]. Our experiments thus point to a rapid, spatially distributed gain control mechanism that emerges in early visual processing.

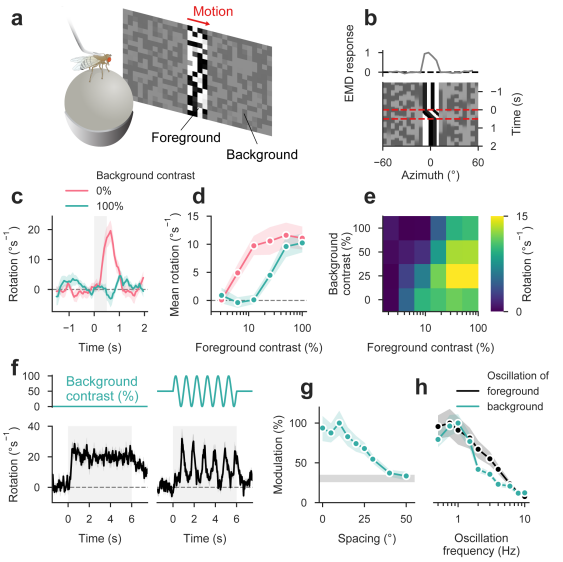


Figure 2. Sensitivity of Drosophila optomotor response is controlled by surround contrast. a, Experimental set-up. Visual display is separated into two areas whose contrast can be set independently. **b,** Bottom, space-time plot of base stimulus. Foreground pattern moved during time span indicated by 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). Gray-shaded area indicates motion. **d,** Mean rotation (averaged between 0 and 1s after stimulus onset) as function of foreground contrast for two background conditions (N=16). **e**, Heatmap of mean rotation for multiple background conditions. With increasing background contrast optomotor sensitivity shifted rightward (N=16). **f,** Stimulus for mapping magnitude of sensitivity shift (N=16). Left, baseline turning response in the absence of background contrast (foreground contrast 25%). Right, turning response for sinusoidal change in background contrast. During high-contrast phase, optomotor response was suppressed; turning modulation allowed read-out of background-induced changes in gain. **g,** Turning response modulation as a function of spacing between motion stimulus and background (N=16). Gray-shaded bar indicates 68% confidence interval around baseline modulation in the absence of background. **h,** 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.

We next used two-photon calcium imaging to locate the neural origin of contrast adaptation. Fly motion vision circuitry is arranged in cartridges processing visual input retinotopically. In various combinations, lamina cells L1–5 feed into at least four medulla cell types per ON and OFF pathway[3, 21]. Medulla units fall into two classes characterized either by transient temporal filtering and moderate center-surround antagonism in their spatial receptive field (Mi1, 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)[22]. T4 and T5 cells then compute direction by comparing medulla signals with different dynamics across neighboring cartridges[3, 9, 21]. Here, we additionally estimated linear receptive fields for L1–5 using stochastic stimuli as before[22]

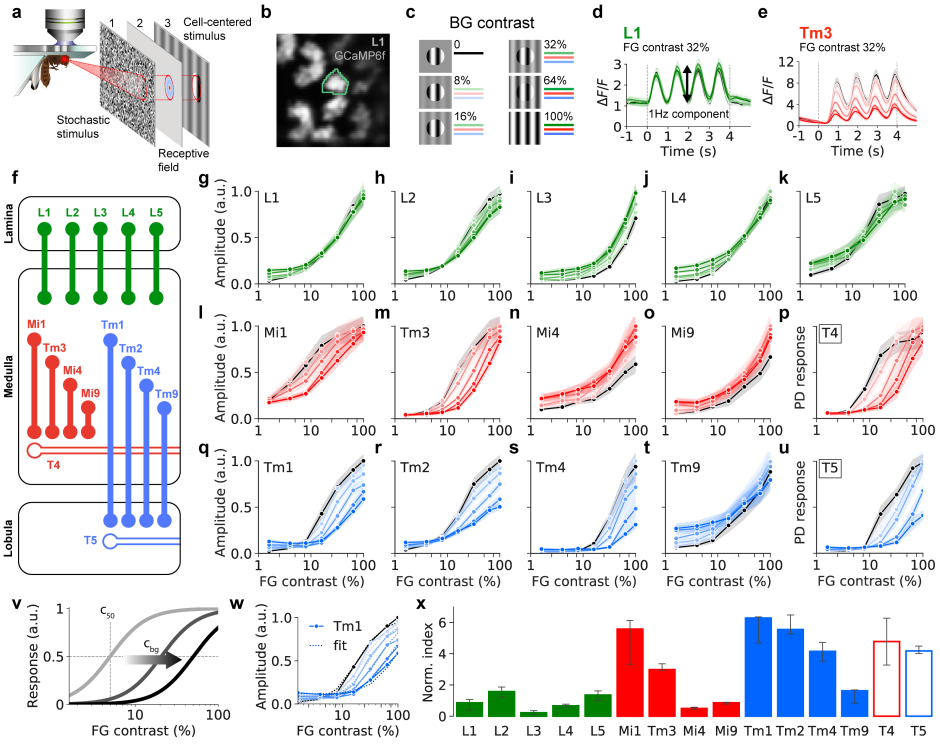


Figure 3. Contrast normalization emerges in transient medulla neurons. **a,** Schematic of experimental procedure: 1. White noise stimulus. 2. Receptive field reconstruction from single neuron calcium signals. 3. Drifting grating with different contrasts in foreground and background. **b,** Two-photon 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 panels g–u. Zero background contrast condition is shown in black. **d,e,** Average calcium responses of L1 and Tm3 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: 21/7 cells/flies, L2: 26/8, L3: 23/6, L4: 19/6, L5: 18/9). **l–p,** Contrast tuning curves for ON pathway neurons (Mi1: 20/5, Tm3: 21/8, Mi4: 20/13, Mi9: 21/9, T4: 23/10). **q–u,** Contrast tuning curves for OFF pathway neurons. (Tm1: 21/7, Tm2: 20/6, Tm4: 20/13, Tm9: 19/6, T5: 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/T5 exhibited strongest degree of normalization.

(Fig. S3a–t). Consistent with previous functional work[13, 23], spatiotemporal filters grouped into tonic (L3) or transient units (L1, L2, L4, and L5) like in the medulla.

To precisely map context-dependent changes in contrast sensitivity for these cell types, we targeted visual stimuli to individual neurons by determining receptive field coordinates of single axon terminals through a combination of stochastic stimuli and online reverse-correlation (Fig. 3a,b; see Methods). Analogously to behavioral experiments, we then presented drifting sine gratings with independently controlled contrast in foreground (as defined by a 25° circular window centered on the receptive field) and background (Fig. 3c). At fixed foreground contrast, L1 activity followed local grating luminance regardless of background (Fig. 3d). Responses in downstream synaptic partner Tm3, however, showed the signature of gain control as signal amplitude was increasingly suppressed by growing surround contrast (Fig. 3e).

We evaluated calcium modulation at the stimulus frequency to obtain contrast response curves for all columnar cell types and T4/T5 cells (Fig. 3f). Lamina units tracked foreground contrast approximately linearly but were weakly if at all modulated by the surround except for a vertical shift at low levels, likely due to background leaking into the receptive fields (Fig. 3g–k, Fig. S4). In the medulla (Fig. 3l–u), tonic Mi4, Mi9, and Tm9 showed similar tuning as L1–5 and again little surround-dependency. For all transient cells (Mi1 and Tm3 for ON; Tm1, Tm2, and Tm4 for OFF), however, tuning deviated from linearity and increasing background contrast had a strongly suppressive effect. Curves were shifted rightward on the logarithmic axis which corresponds to divisive stretching in linear contrast space. Sensitivity to foreground contrast was generally higher in ON than in OFF units. Importantly, preferred direction responses in T4 and T5 were strongly background-dependent (Fig. 3p,u) even though not all their medulla inputs are subject to gain control.

To quantify tuning curves in detail, we fit a closed-form model based on a common neural computation, divisive normalization (Fig. 3v; see Methods). Here, responses are effectively divided by an inhibitory signal that estimates surround contrast as the average activity of a pool of nearby neurons[6, 7]. Divisive normalization thus compresses signals of varying contrast into a fixed range by dynamically adjusting gain to current conditions[1, 6]. The model accurately reproduced tuning curves for each cell type (Fig. 3w, Table S2).

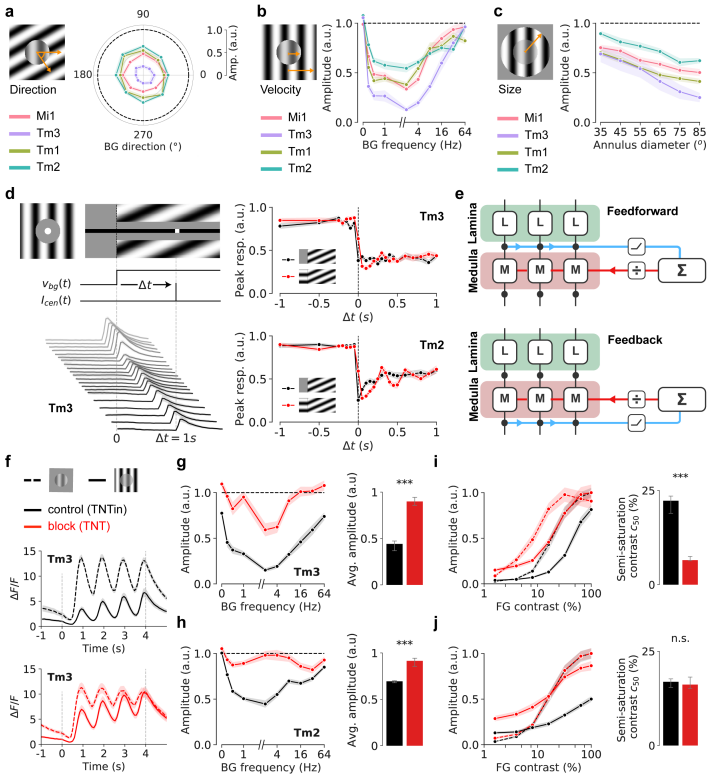


Figure 4. Neural contrast normalization relies on rapid feedback from a pool of transient units. **a,** Polar plot of response amplitude for different directions of background motion. Black dashed line represents reference measured with background contrast of 0%. For each neuron, foreground contrast was chosen to maximize possible background suppression (Mi1: 16%, Tm3: 32%, Tm1: 64%, Tm2: 100%). **b,** Responses for different background contrast frequencies, revealing transient tuning of suppression. **c,** Suppression strength increased with outer diameter of background annulus (Mi1: 21/9 cells/flies, 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). **e,** Schematic of feedforward and feedback model for surround suppression. **f,** Mean responses of Tm3 for TNT block (red) and TNTin controls at background frequency 16 Hz (black; dashed line indicates reference response and solid line the response at full background contrast). **g,** Left, frequency tuning for block experiments (as in b). Right, average amplitude over all frequencies was higher for Tm3 block flies (Tm3 block: 21/5, Tm3 control: 20/5; Mann-Whitney U: 8, ***P<0.001). **h,** Blocking results for Tm2 (as in g; Tm2 block: 20/5, Tm2 control: 25/6; Mann-Whitney U: 17, ***P<0.001). **i,** Left, foreground contrast tuning for block experiments at 0 and 100% background contrast. Right, contrast sensitivity as measured by semi-saturation constant c_{50} was increased for Tm3 block flies (Mann-Whitney U: 39, ***P<0.001). **j,** Blocking results for Tm2 (as in i; 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.

Critically, it accounted for vertical shifts through linear background leakage as well as sigmoidal tuning curves and context-dependent changes in contrast sensitivity through the divisive term. We computed a normalization index from estimated model parameters that quantified the factor by which tuning curves would shift when background contrast was increased from 0% to 100% (see Methods). The index was substantially higher in transient medulla cells (Mi1, Tm3, Tm1, Tm2, and Tm4) and direction-selective T4/T5 cells than in L1–5 or tonic medulla units (Mi4, Mi9, and Tm9; Fig. 3x). Curiously, L2 and L5 exhibited mildly elevated normalization indices which may be due to particular connectivity with parallel neurons[3] or the proposed complex receptive field structure of L2[24].

Thus, 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 contrast (Fig. 4a). Contrast frequency tunings for suppression resembled band-pass filters with a peak at 2 Hz (Fig. 4b). Crucially, static backgrounds did not have a suppressive effect. Both isotropy and frequency tunings were strikingly similar to filter properties of transient lamina and medulla units (Fig. S3u,v). Suppression steadily increased with the outer diameter of an annulus containing the background pattern, which again indicated an extended integration area (Fig. 4c). Spatiotemporal features of neural gain control thus matched our findings from behavior (Fig. 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's polarity (Fig. 4d). By varying time between background grating onset and pulse, we scanned the temporal profile of the suppressive signal. For the tested neurons Tm3 and Tm2, we found virtu-

ally immediate response reduction within a measurement precision of 50ms. We observed transient ringing of suppression strength at the background frequency. Ringing was stronger when the grating was present before motion onset compared to when it was masked. A similar effect has been described in LPTCs[11] where it results from neural integration of multiple transient, out-of-phase inputs. In sum, these findings suggest that surround suppression derives from a pool of direction-agnostic transient neurons. We built a time-resolved, data-driven model to determine whether a mechanism that pools transient units across space to divisively suppress local responses could reproduce our findings. The model faithfully predicted direction, frequency, and size tunings as well as contrast-step ringing, T4/T5 responses, and LPTC output for our behavioral stimuli (Fig. S5a–k).

Spatial pooling, however, could occur over either feedforward signals from the lamina or feedback from the medulla (Fig. 4e). In vertebrate systems it has proven difficult to distinguish the two[6]. Fly transient units in lamina or medulla have similar temporal properties (Fig. S3u,v) and both implementations produce equivalent steady-state output[7], so we used genetic silencing to pinpoint the source. We co-expressed a calcium indicator and the tetanus toxin light chain (TNT; see Methods) in different cell types, blocking synaptic output and thus feedback from the entire neuron array but leaving feedforward input and calcium signals intact. Compared to controls with inactive TNT, we observed significantly reduced suppression across background frequencies for ON cells Mi1 and Tm3 as well as OFF-sensitive Tm2 but not Tm1 (Fig. 4f,g,h and Fig. S6a,b). In Mi1 and Tm3, baseline contrast sensitivity as measured by the semi-saturation constant of model fits was significantly increased (Fig. 4i and Fig. S6c), suggesting that cells were disinhibited due to reduced pool signal. For Tm2, linear background contribution at low foreground contrasts increased whereas high contrast responses were less suppressed by the surround (Fig. 4j). In the fly, contrast normalization is therefore at least partially based on feedback from a combination of medulla neurons.

Does this type of response normalization account for the robustness of fly motion detection? Previous work on EMDs and natural scenes has exploited compressive transforms but did so heuristically or without surround-dependent gain control[15, 16]. We evaluated natural image responses in the data-driven LPTC model and found moderate reduction of cross-image variability compared to a model with bypassed normalization (Fig. S5l–n). 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[8, 25]. EMD-like architectures are concisely expressed as multi-layer convolutional networks[8] and fully differentiable, rendering them amenable to optimization methods like gradient descent. We designed a fly-like neural network and independently trained possible types of contrast processing such that each model class could optimally adapt

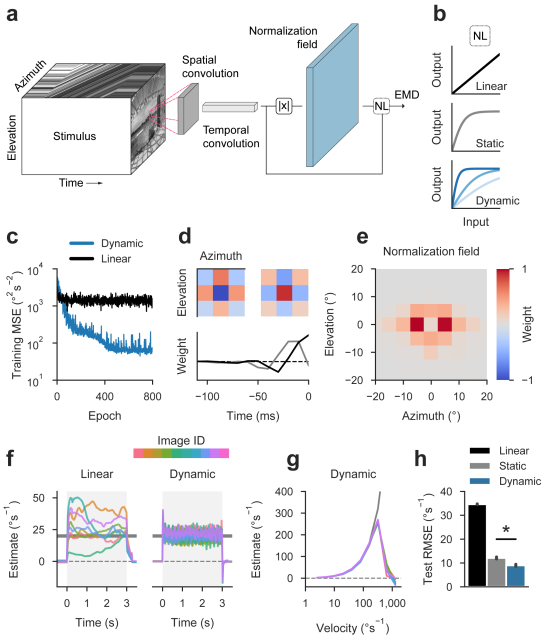


Figure 5. Contrast normalization enhances robustness to natural scene variability. a, Schematic of single convolutional input filter. Motion stimuli are sequentially processed by a spatial 3 x 3 x 1 (azimuth, elevation, time) and a temporal 1 x 1 x 30 filter. Through a transfer function, the signal is combined with a normalization signal generated by a 11 x 11 x 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 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°s^{-1} 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 3s after motion onset). Gray line indicates true velocity. **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.

to a specific, behaviorally relevant task.

All models featured linear, spatiotemporally separable input convolutions (Fig. 5a). 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 (Fig. 5a,b; see 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 (Fig. 5c). 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 (Fig. 5d,

Fig. S7a–c). Models thus made extensive use of redundancy reduction through center-surround configurations[20] and discovered the EMD strategy of delay-and-compare[11]. Normalization fields for the dynamic model spanned approximately 30° in azimuth and invariably excluded information from the center of the filter (Fig. 5e). Interestingly, dynamic models exploited normalization in both channels and switched normalization strategies during training, transitioning from purely static to purely context-dependent compression (Fig. S7d,e). Overall, normalized networks acquired representations that matched filtering and gain control properties of the fly medulla.

When tested on previous experimental stimuli (Fig. 1), linear models exhibited improved velocity tuning curves compared to a standard EMD (Fig. 1f and Fig. 5f, Fig. S5f) 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 (Fig. 5f,g). 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 (Fig. 5h). 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 (Fig. S7g) where linear models failed catastrophically while both compressive stages retained performance. This was particularly pronounced when testing images with high dynamic range (see Methods). Critically, on all datasets dynamic compression resulted in substantial error reduction with respect to both linear transfer and static compression.

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. Normalization has been described as a generic mechanism for removing higher-order correlations from natural signals[1, 26]. Here, we close the loop between neural mechanism and an ecologically critical behavior, the optomotor response, and demonstrate how contrast gain control renders motion detection resilient to challenges imposed by natural scene statistics. Our convolutional network solves the task of estimating velocity across diverse environments and at little computational cost, particularly compared to standard optic flow algorithms. Present findings may thus aid the design of low-power, low-latency machine vision systems suitable for autonomous vehicles[27]. Gain control in the Drosophila optic lobe bears striking resemblance to normalization in other modalities like the fly olfactory system[28] or mammalian auditory cortex[29] as well as processing in vertebrate visual areas from retina to V1[6]. The tuning of non-linear surround suppression in lateral geniculate nucleus, in particular, matches that of transient units in the fly medulla[30]. This provides further proof for evolutionary convergence on canonical algorithms[6].

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 modelling, and wrote the manuscript. All authors participated in editing the manuscript.

Author Information. The authors declare no competing interests. Correspondence and requests for materials should be addressed to A.L. (leonhardt@neuro.mpg.de) or M.S.D. (drews@neuro.mpg.de).

ACKNOWLEDGEMENTS
We thank S. Prech for technical assistance, J. Kuhl for schematics, A. Nern 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.

References

1. F. Rieke and M. E. Rudd, "The challenges natural images pose for visual adaptation," *Neuron*, vol. 64, no. 5, pp. 605–616, 2009.
2. W. S. Geisler, "Visual perception and the statistical properties of natural scenes," *Annual Review of Psychology*, vol. 59, pp. 167–192, 2008.
3. A. S. Mauss, A. Vlasits, A. Borst, and M. Feller, "Visual circuits for direction selectivity," *Annual Review of Neuroscience*, vol. 40, pp. 211–230, July 2017.
4. A. Leonhardt, G. Ammer, M. Meier, E. Serbe, A. Bahl, and A. Borst, "Asymmetry of *Drosophila* ON and OFF motion detectors enhances real-world velocity estimation," *Nature Neuroscience*, vol. 19, pp. 706–715, May 2016.
5. A. D. Straw, T. Rainsford, and D. C. O'Carroll, "Contrast sensitivity of insect motion detectors to natural images," *Journal of Vision*, vol. 8, no. 3, pp. 32–32, 2008.
6. M. Carandini and D. J. Heeger, "Normalization as a canonical neural computation," *Nature Reviews Neuroscience*, vol. 13, pp. 51–62, Nov. 2011.
7. D. J. Heeger, "Normalization of cell responses in cat striate cortex," *Visual neuroscience*, vol. 9, pp. 181–197, Aug. 1992.
8. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
9. M. S. Maisak, J. Haag, G. Ammer, E. Serbe, M. Meier, A. Leonhardt, T. Schilling, A. Bahl, G. M. Rubin, A. Nern, B. J. Dickson, D. F. Reiff, E. Hopp, and A. Borst, "A directional tuning map of *Drosophila* elementary motion detectors," *Nature*, vol. 500, pp. 212–216, Aug. 2013.
10. A. Bahl, E. Serbe, M. Meier, G. Ammer, and A. Borst, "Neural mechanisms for *Drosophila* contrast vision," *Neuron*, vol. 88, no. 6, pp. 1240–1252, 2015.
11. A. Borst, C. Reisenman, and J. Haag, "Adaptation of response transients in fly motion vision. II: Model studies," *Vision Research*, vol. 43, pp. 1311–1324, May 2003.
12. A. Borst, V. L. Flanagin, and H. Sompolinsky, "Adaptation without parameter change: dynamic gain control in motion detection," *Proceedings of the National Academy of Sciences* , vol. 102, pp. 6172–6176, Apr. 2005.
13. D. A. Clark, L. Bursztyn, M. A. Horowitz, M. J. Schnitzer, and T. R. Clandinin, "Defining the computational structure of the motion detector in *Drosophila*," *Neuron*, vol. 70, pp. 1165–1177, June 2011.
14. J. E. Fitzgerald and D. A. Clark, "Nonlinear circuits for naturalistic visual motion estimation," *Elife*, vol. 4, p. e09123, Oct. 2015.
15. R. O. Dror, D. C. O'Carroll, and S. B. Laughlin, "Accuracy of velocity estimation by reichardt correlators," *Journal of the Optical Society of America. A, Optics and Image Science*, vol. 18, pp. 241–252, Feb. 2001.
16. R. S. A. Brinkworth and D. C. O'Carroll, "Robust models for optic flow coding in natural scenes inspired by insect biology," *PLoS computational biology*, vol. 5, p. e1000555, Nov. 2009.
17. R. A. Harris, D. C. O'Carroll, and S. B. Laughlin, "Contrast gain reduction in fly motion adaptation," *Neuron*, vol. 28, pp. 595–606, Nov. 2000.
18. W. Reichardt, T. Poggio, and K. Hausen, "Figure-ground discrimination by relative movement in the visual system of the fly," *Biological Cybernetics*, vol. 46, pp. 1–30, Jan. 1983.
19. J. H. van Hateren, "Processing of natural time series of intensities by the visual system of the blowfly," *Vision Research*, vol. 37, pp. 3407–3416, Dec. 1997.
20. M. V. Srinivasan, S. B. Laughlin, and A. Dubs, "Predictive coding: a fresh view of inhibition in the retina," *Proceedings of the Royal Society of London. Series B. Biological Sciences*, vol. 216, pp. 427–459, Nov. 1982.
21. K. Shinomiya, G. Huang, Z. Lu, T. Parag, C. S. Xu, R. Aniceto, N. Ansari, N. Cheatham, S. Lauchie, E. Neace, O. Ogundeyi, C. Ordish, D. Peel, A. Shinomiya, C. Smith, S. Takemura, I. Talebi, P. K. Rivlin, A. Nern, L. K. Scheffer, S. M. Plaza, and I. A. Meinertzhagen, "Comparisons between the ON- and OFF-edge motion pathways in the *Drosophila* brain," *eLife*, vol. 8, p. e40025, 2019.

22. A. Arenz, M. S. Drews, F. G. Richter, G. Ammer, and A. Borst, "The temporal tuning of the *Drosophila* motion detectors is determined by the dynamics of their input elements," *Current Biology*, vol. 27, pp. 929–944, Apr. 2017.
23. M. Silies, D. M. Gohl, Y. E. Fisher, L. Freifeld, D. A. Clark, and T. R. Clandinin, "Modular use of peripheral input channels tunes motion-detecting circuitry," *Neuron*, vol. 79, pp. 111–127, July 2013.
24. L. Freifeld, D. A. Clark, M. J. Schnitzer, M. A. Horowitz, and T. R. Clandinin, "GABAergic lateral interactions tune the early stages of visual processing in *Drosophila*," *Neuron*, vol. 78, pp. 1075–1089, June 2013.
25. D. L. Yamins and J. J. DiCarlo, "Using goal-driven deep learning models to understand sensory cortex," *Nature neuroscience*, vol. 19, no. 3, p. 356, 2016.
26. O. Schwartz and E. P. Simoncelli, "Natural signal statistics and sensory gain control," *Nature Neuroscience*, vol. 4, pp. 819–825, Aug. 2001.
27. J. Plett, A. Bahl, M. Buss, K. Kühnlenz, and A. Borst, "Bio-inspired visual ego-rotation sensor for mavs," *Biological cybernetics*, vol. 106, no. 1, pp. 51–63, 2012.
28. S. R. Olsen, V. Bhandawat, and R. I. Wilson, "Divisive normalization in olfactory population codes," *Neuron*, vol. 66, pp. 287–299, Apr. 2010.
29. N. C. Rabinowitz, B. D. B. Willmore, J. W. H. Schnupp, and A. J. King, "Contrast gain control in auditory cortex," *Neuron*, vol. 70, pp. 1178–1191, June 2011.
30. V. Bonin, V. Mante, and M. Carandini, "The suppressive field of neurons in lateral geniculate nucleus," *Journal of Neuroscience*, vol. 25, pp. 10844–10856, Nov. 2005.

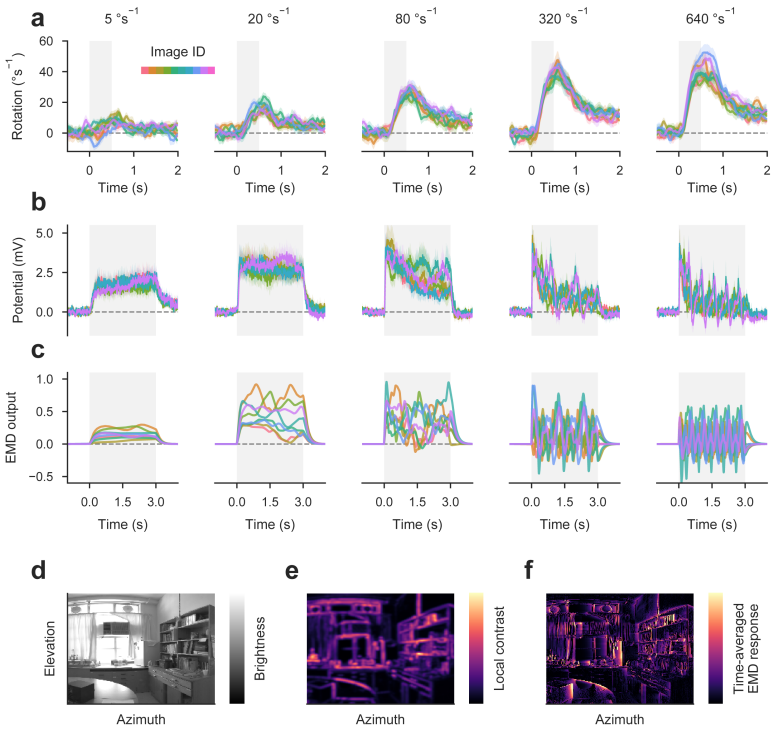


Figure S1. Behavioral, neural, and model responses to natural scenes (related to Fig. 1). a, Turning responses for 8 images (indicated by trace color) and 5 velocities (indicated by panel title; N=16 wild-type flies; data as in Fig. 1b). Gray shaded area indicates duration of motion stimulus. **a,** Membrane potential for 6 images and 5 velocities (N=11 HS cells from 9 flies; data as in Fig. 1d). **c,** Output of model LPTC for same images and velocities as b (data as in Fig. 1f). **d,** Natural image patch as seen through the field of view of model LPTC. **e,** 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. **f,** 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 c (see Methods) with the exception of more fine-grained sampling at exactly the image resolution. The depicted panorama was moved for 16s at a velocity of $22.5^\circ s^{-1}$, 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.

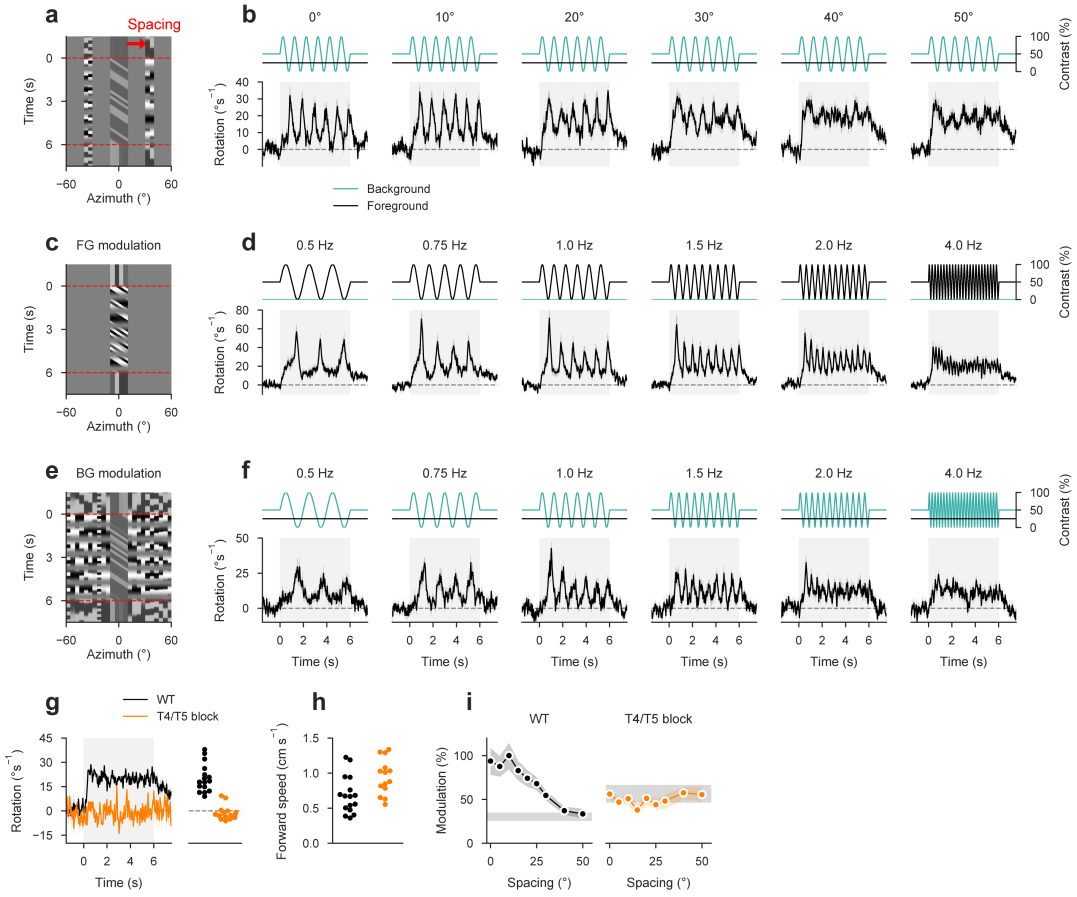


Figure S2. Detailed behavioral responses to contrast stimuli (related to Fig. 2). a, Illustration of spatial oscillation experiment. Background was restricted to 10° wide stripes flanking the foreground motion stimulus at the spacing indicated by red arrow. Dashed lines indicate period during which foreground 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 spacing 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,** Left, comparison of turning responses between wild-type flies and flies in which T4/T5 cells were silenced using TNT (see Methods; N=16/14 for WT/block flies). Right, turning responses averaged between 0 and 6s following motion onset. Syndirectional turning was abolished in T4/T5-silenced flies. **h,** Average forward speed throughout full experiment. T4/T5 block flies did not exhibit locomotion deficiencies. **i,** 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.

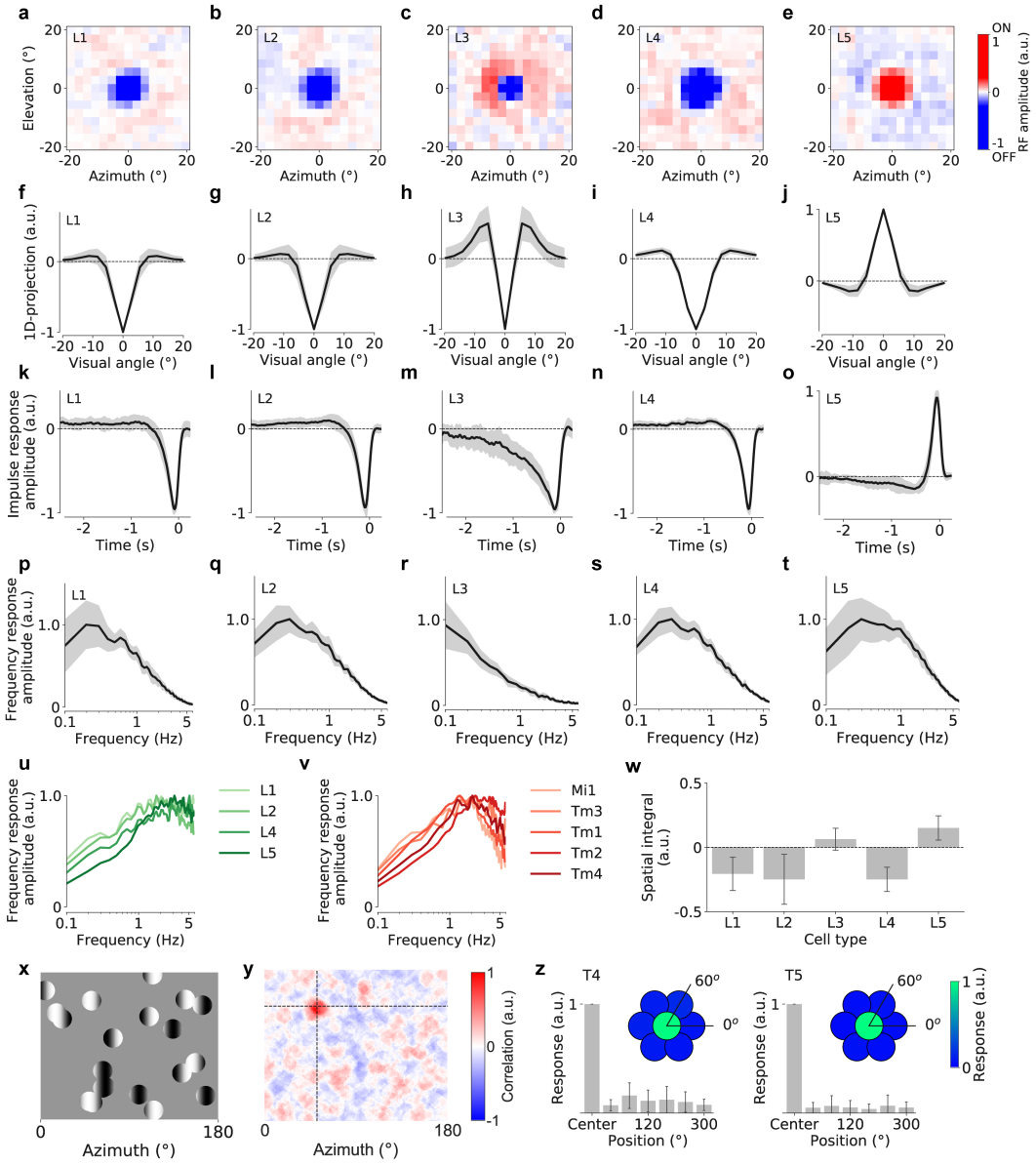


Figure S3. Lamina and T4/T5 receptive field mapping. a–e, Averaged 2D spatial receptive fields (RF) of L1–L5 from reverse correlation using white noise stimulation (L1: 21/7 cells/flies, L2: 34/5, L3: 34/5, L4: 17/6, L5: 18/9). **f–j,** 1D projection (averaged over all orientations) of the RFs in a–e. 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/band-pass cells (all lamina cells except for L3) after deconvolution with a putative linear GCaMP6f low-pass filter with time constant 350ms as performed previously[22]. **v,** Deconvolved frequency responses of medulla bandpass filter cells (replotted from previous work[22]). **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.

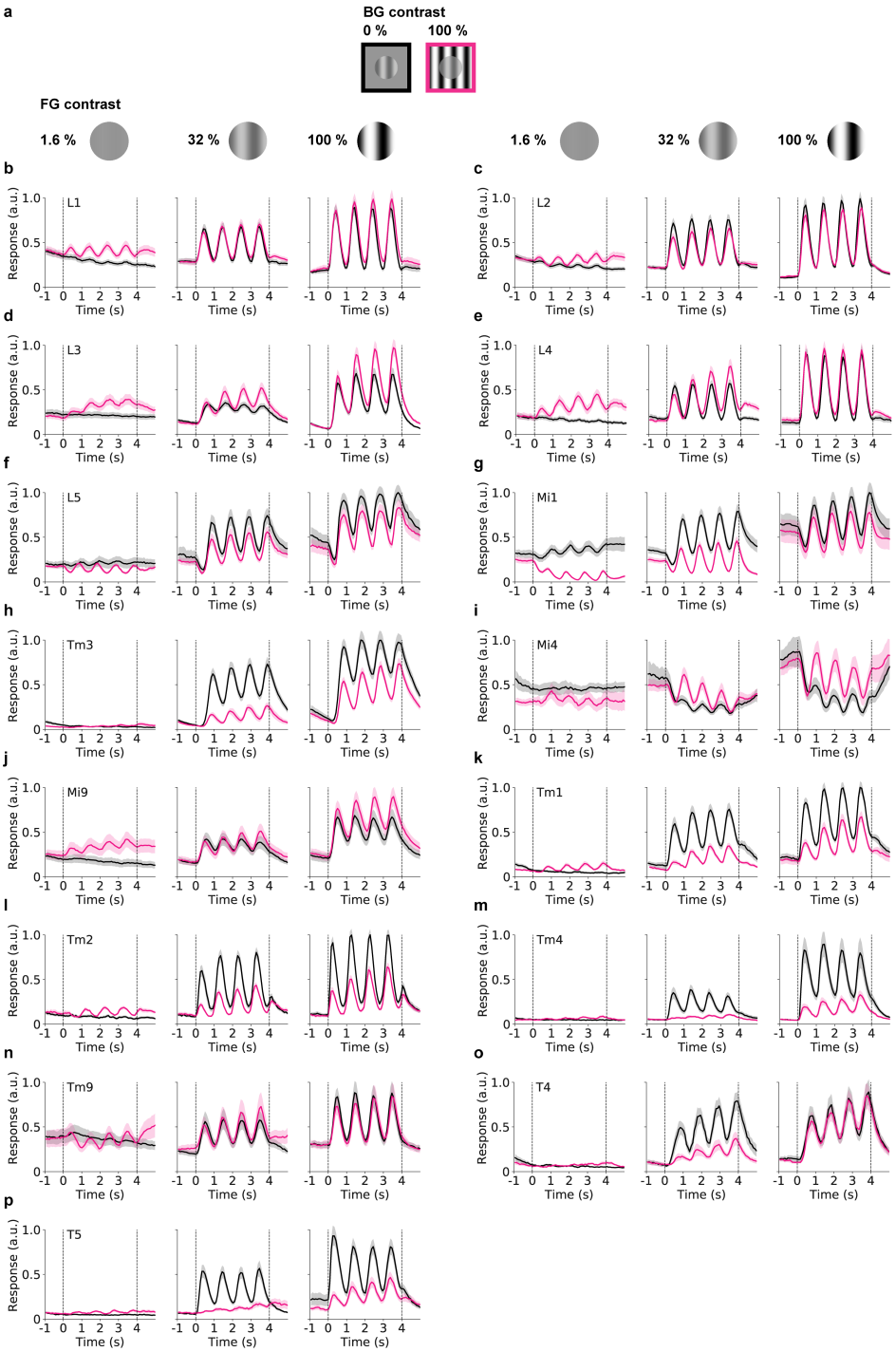


Figure S4. Raw calcium responses for basic contrast stimuli (related to Fig. 3). a, Shown is only a subset of the data evaluated in Fig. 3. Background contrast of 0% is indicated by black lines, background contrast of 100% is depicted in magenta. Responses are shown only for 3 out of 7 foreground contrasts. **b–p,** Average calcium responses of all neurons to combinations of different foreground and background contrasts.

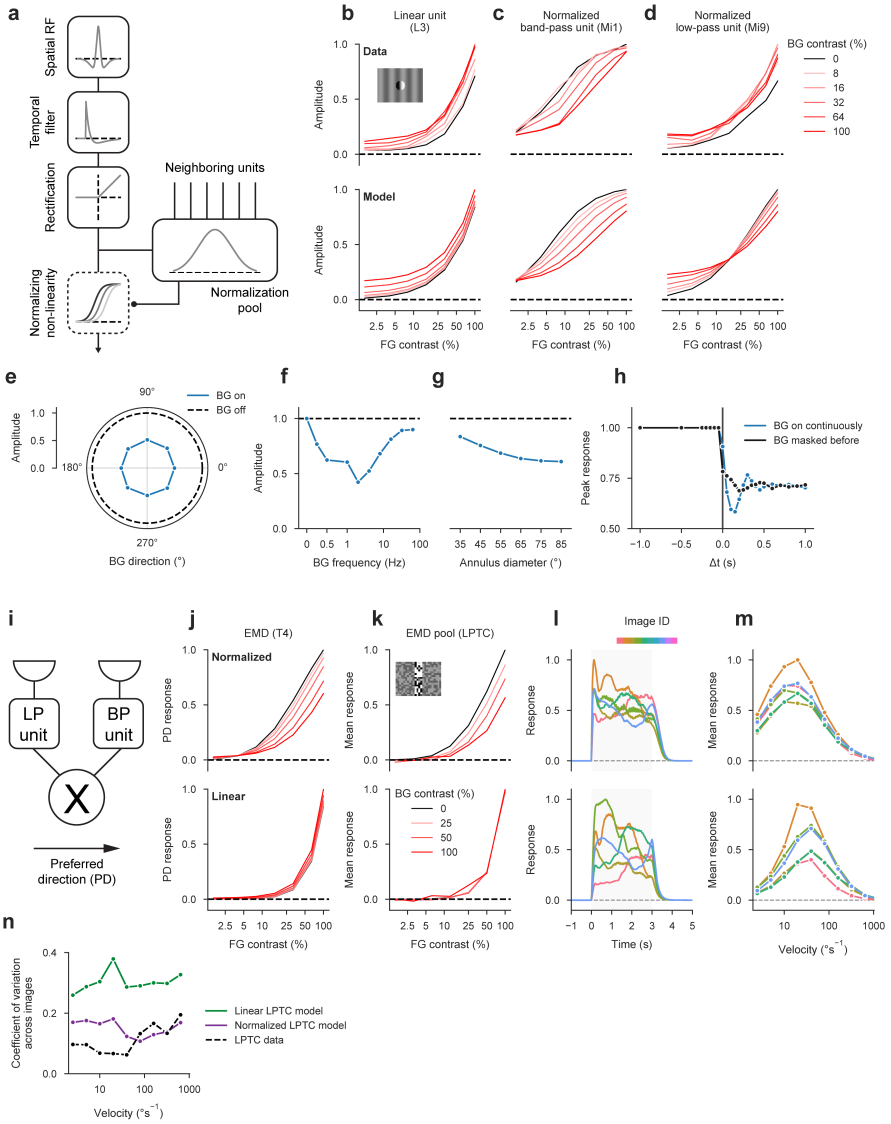


Figure S5. Data-driven functional model of normalization circuit. a, Illustration of signal cascade for data-driven cell model (see Methods for details). Filter elements are sketched for an ON band-pass cell with normalization. **b–d,** Contrast tuning curves for three model cells, estimated using the same protocol as during calcium imaging (FG = foreground, BG = background). Top, empirical data for L3, Mi1, and Mi9 (see Fig. 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 Methods for parameters). **e,** Responses of normalized ON band-pass cell model to orientation tuning stimulus (see Fig. 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 Fig. 4b; dashed line marks reference stimulus without background). **g,** Responses of the same model to background size stimulus (see Fig. 4c; dashed line marks reference stimulus without background). **h,** Responses of the same model to contrast-step protocol (see Fig. 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 Fig. 3). **k,** Responses to behavioral contrast stimulus for a LPTC model composed of T4 and T5 models (see Methods). **l,** Responses to various natural scenes moving at $20^\circ s^{-1}$ (modelled and evaluated as in Fig. 1). **m,** Velocity tuning curves for natural scenes (modelled and evaluated as in Fig. 1). **n,** Coefficient of variation across images for individual image velocities (derived from velocity tuning curves in m and Fig. 1f; see Methods for details). A model including input normalization outperformed the linear model and approximated the variability of LPTC responses.

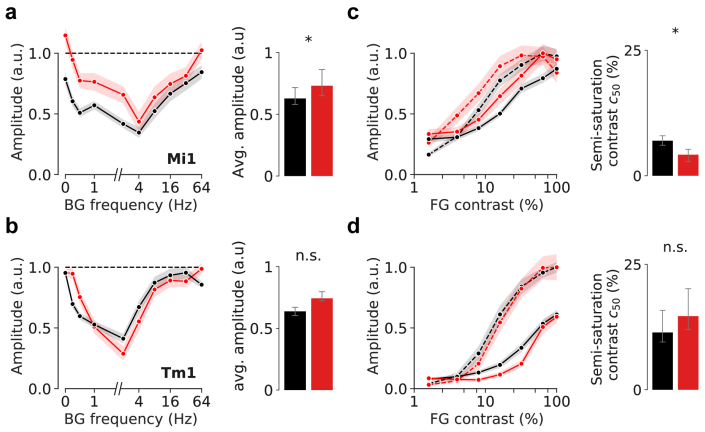


Figure S6. Additional silencing data for contrast tuning experiment (related to Fig. 4). a, Left, frequency tuning of suppression as in Fig. 4b,g,h. Black dashed line represents reference stimulus with background contrast of 0%. Right, average amplitude over all frequencies was higher for Mi1 block flies which indicates reduced suppression (Mi1 block: 20/5 cells/flies, Mi1 control: 21/6; Mann-Whitney U: 143, *P=0.04). **b,** Tm1 data as in a. There was no significant effect when blocking Tm1 cells (Tm1 block: 20/5, Tm1 control: 19/5; Mann-Whitney U: 169, NS P=0.28). **c,** Left, foreground contrast tuning for 0 (dashed line) and 100% (solid line) background contrast as in Fig. 4i,j. Right, contrast sensitivity as measured by semi-saturation constant c_{50} of model fit to data was increased for Mi1 block flies (Mann-Whitney U: 128, *P=0.02). **d,** Tm1 data as in c. There was no significant effect when blocking Tm1 cells (Mann-Whitney U: 158, NS P=0.19). Error bars show bootstrapped 68% confidence intervals around the median.

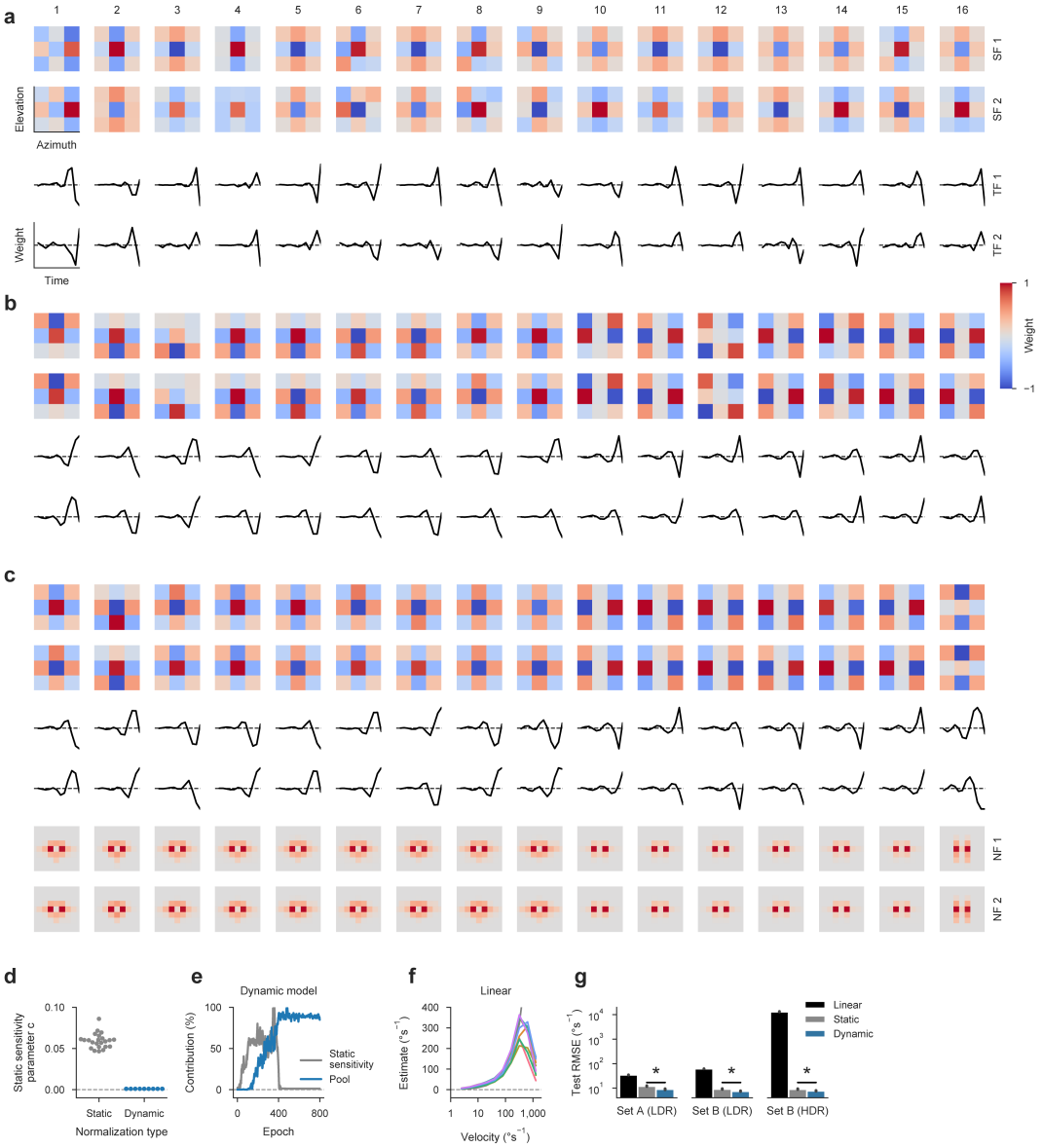


Figure S7. Detailed receptive fields and performance data for task-driven model (related to Fig. 5). a–c, Receptive fields and temporal filters for 16 models of each non-linearity configuration (a, linear; b, static; c, dynamic). Models were sorted by test set error (increasing from left to right). Each 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 Fig. 5. **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 Fig. 5g). Gray curve indicates true scene velocity on logarithmic axis. **g,** Quantification of average model performance for all tested data sets (analogously to Fig. 5h; LDR = low dynamic range, HDR = high dynamic range). See 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.

Short name	Full genotype	Used in
WT	w+/w+; +/+; +/+	Fig. 1, Fig. 2
T4/T5 block	w+/w-; R59E08-AD/UAS-TNT; R42F06-DBD/+	Fig. S2
L1-GCaMP6f	w+/w-; VT027316-AD/UAS-GCaMP6f; R40F12-DBD/UAS-GCaMP6f	Fig. 3, Fig. S3, Fig. S4
L2-GCaMP6f	w+/w-; R53G02-AD/UAS-GCaMP6f; R29G11-DBD/UAS-GCaMP6f	Fig. 3, Fig. S3, Fig. S4
L3-GCaMP6f	w+/w-; R59A05-AD/UAS-GCaMP6f; R57H07-DBD/UAS-GCaMP6f	Fig. 3, Fig. S3, Fig. S4
L4-GCaMP6f	w+/w-; R20A03-AD/UAS-GCaMP6f; R31C06-DBD/UAS-GCaMP6f	Fig. 3, Fig. S3, Fig. S4
L5-GCaMP6f	w+/w-; R21A05-AD/UAS-GCaMP6f; R31H09-DBD/UAS-GCaMP6f	Fig. 3, Fig. S3, Fig. S4
Mi1-GCaMP6f	w+/w-; R19F01-AD/UAS-GCaMP6f; R71D01-DBD/UAS-GCaMP6f	Fig. 3, Fig. S4, Fig. S4
Mi4-GCaMP6f	w+/w-; R48A07-AD/UAS-GCaMP6f; R13F11-DBD/UAS-GCaMP6f	Fig. 3, Fig. S4
Mi9-GCaMP6f	w+/w-; R48A07-AD/UAS-GCaMP6f; VT046779-DBD/UAS-GCaMP6f	Fig. 3, Fig. S4
Tm1-GCaMP6f	w+/w-; R41G07-AD/UAS-GCaMP6f; R47G01-DBD/UAS-GCaMP6f	Fig. 3, Fig. S4, Fig. S4
Tm2-GCaMP6f	w+/w-; +/UAS-GCaMP6f; VT12282/UAS-GCaMP6f	Fig. 3, Fig. 4a-c, Fig. S4
Tm2split-GCaMP6f	w+/w-; R28D05-AD/UAS-GCaMP6f; R82F12-DBD/UAS-GCaMP6f	Fig. 4d
Tm4-GCaMP6f	w+/w-; +/UAS-GCaMP6f; R35H01/UAS-GCaMP6f	Fig. 3, Fig. S4
Tm9-GCaMP6f	w+/w-; +/UAS-GCaMP6f; VT65303/UAS-GCaMP6f	Fig. 3, Fig. S4
T4-GCaMP6f	w+/w-; VT16255-AD/UAS-GCaMP6f; VT12314-DBD/UAS-GCaMP6f	Fig. 3, Fig. S3, Fig. S4
T5-GCaMP6f	w+/w-; VT13975-AD/UAS-GCaMP6f; R42H07-DBD/UAS-GCaMP6f	Fig. 3, Fig. S3, Fig. S4
Mi1-GCaMP6f, TNT-E	w+/w-; R19F01-AD/UAS-TNT-E; R71D01-DBD/UAS-GCaMP6f	Fig. S6
Mi1-GCaMP6f, TNTin	w+/w-; R19F01-AD/UAS-TNTin; R71D01-DBD/UAS-GCaMP6f	Fig. S6
Tm3-GCaMP6f, TNT-E	w+/w-; R13E12-AD/UAS-TNT-E; R59C10-DBD/UAS-GCaMP6f	Fig. 4
Tm3-GCaMP6f, TNTin	w+/w-; R13E12-AD/UAS-TNTin; R59C10-DBD/UAS-GCaMP6f	Fig. 4
Tm1-GCaMP6f, TNT-E	w+/w-; R41G07-AD/UAS-TNT-E; R47G01-DBD/UAS-GCaMP6f	Fig. S6
Tm1-GCaMP6f, TNTin	w+/w-; R41G07-AD/UAS-TNTin; R47G01-DBD/UAS-GCaMP6f	Fig. S6
Tm2split-GCaMP6f, TNT-E	w+/w-; R28D05-AD/UAS-TNT-E; R82F12-DBD/UAS-GCaMP6f	Fig. 4
Tm2split-GCaMP6f, TNTin	w+/w-; R28D05-AD/UAS-TNTin; R82F12-DBD/UAS-GCaMP6f	Fig. 4

Table S1. Genotypes and abbreviations

Cell type	$\mathbf{L_{fg}}$	$\mathbf{L_{bg}}$	$\mathbf{p}$	$\mathbf{c_{50}}$	$\mathbf{w_{pool}}$	$\mathbf{q}$	**Norm. index**	$\mathbf{R^2_{DivisiveNorm}}$	$\mathbf{R^2_{linear}}$
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 S2. Fits for divisive normalization model (related to Fig. 3)

Methods

Flies and genetics. 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[31]. Resulting genotypes and their shorthands are listed in Table S1.
Unless stated otherwise, locomotion and tangential cell responses were recorded in wild-type Canton S flies 1 to 5 days after eclosion (Fig. 1 and Fig. 2). We used the genetically encoded calcium indicator GCaMP6f[32] to determine the functional properties of individual cell types (Fig. 3 and Fig. 4). Throughout silencing experiments (Fig. 4, Fig. S2 and Fig. S5), we expressed tetanus toxin light chain (TNT) or an inactive version (TNTin) in the cell type of interest[33]. For calcium imaging experiments involving silencing (Fig. 4 and Fig. S6), one day old flies were collected and put on 29°C for 3 days to boost expression of TNT or TNTin.

Natural image sets. For electrophysiology, behavioral, and modelling experiments, we used images from a published set of 20 natural panoramic scenes[16] termed data set A. All images were independently processed as follows: We averaged across color channels and downsampled the scene to a resolution of 1,600 x 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 (Fig. 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 (Fig. 5), 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[34] consisting of 421 images (Fig. S7f). These scenes were kept at their native resolution of 927 x 251 pixels (corresponding to an azimuthal sampling rate of 0.39 pixels per degree) and processed as above, yielding data set 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 data set with high dynamic range (HDR).

Behavioral experiments.

Locomotion recording. Experiments on the treadmill setup were conducted as described before[4, 10, 35]. 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 read-out 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.

Visual stimuli. 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 (Fig. 1) for 8 distinct natural images at 6 logarithmically spaced velocities ranging from 5 to 1,280°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 1.5s, then were rotated at the chosen velocity for 0.5s, and remained fixed for another 1.5s.
The optomotor contrast stimulus separated the visual field into two areas (see Fig. 2a and Fig. S2). For the so-called background, we tiled the visual field with pixels of size 5° x 5°. At each pixel location we drew a temporal frequency f from a normal distribution (μ = 0 Hz, σ = 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

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

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

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

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 Fig. 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° x 5°. For each pixel i, luminance was fixed over time and determined by

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

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.

For the basic contrast tuning experiment (Fig. 2a–e; see Supplementary Video 1), 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 2.0s following stimulus onset, the foreground pattern moved at a fixed velocity of $50°s^{-1}$. For oscillation experiments (Fig. 2f–h), 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 Fig. 2f, Fig. S2a,c,e, and Supplementary Video 2). 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 Fig. S2a and Supplementary Video 3). The spacing (0, 5, 10, 15, 20, 25, 30, 40, or 50°) determined the distance between 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. For the temporal experiments, we measured oscillation frequencies of 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, 8, and 10 Hz (Fig. 2h). 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.

Data evaluation. To ensure data quality, we excluded all flies whose average forward velocity during the experiment was below 0.25 cms^{-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, Fig. 1; 20 Hz for contrast tuning, Fig. 2; 100 Hz for oscillation stimuli, Fig. 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 100ms. 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 analysis.

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

Electrophysiology.

Procedure. Our patch-clamp recordings from tangential cells followed established protocols[36]. 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 stimulus. Visual stimulation was delivered using a cylindrical projector-based arena as previously described[22]. 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±48cd m^{-2} (mean±s.d. 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 $640°s^{-1}$ (Fig. 1). On each presentation, the scene was displayed at a fixed phase, stayed still for 1 s, and then rotated horizontally for 3s 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.

Data evaluation. Voltage data were digitized at 1,000 Hz. To account for slow drift in potential, we subtracted the average voltage in a 1s 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 5mV 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.

Calcium Imaging.

Procedure. Calcium imaging experiments were performed using custom-built two-photon microscopes as described before[22]. The imaging acquisition rate was 11.8 Hz for all experiments, or 23.7 Hz for the experiment in Fig. 4d, with imaging resolutions ranging from 32 x 32 to 64 x 128 pixels. Image acquisition was controlled using the ScanImage software (version 3.8)[37]. We prepared flies as previously described[9, 22]. Briefly, Drosophila were anesthetized on ice and glued onto a plexiglas 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.

Visual stimuli. 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 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/T5 cells, see below). The stimuli were pre-rendered at 60 Hz and generated as previously described[22]. 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 (Fig. S3). This stimulus had a Gaussian autocorrelation with a standard deviation of approximately 45ms in time and a contrast of 25% around a mean intensity value of 50 on an 8 bit grayscale. For some 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/T5, we relied on a novel stochastic motion noise stimulus to determine RF coordinates. First, we determined the preferred direction of a 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 (Fig. S3x). The positions of these 20 windows were changed and randomly chosen every second over 4 min. Reverse correlation of T4/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 (Fig. 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 (Fig. S3z).
For the experiments shown in Fig. 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. 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 seconds after stimulus onset (see Supplementary Video 4). 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 Fig. 4a–c (see Supplementary Videos 5–7), the foreground contrast was chosen depending on the cell type as the point where the suppression elicited by 100% background contrast (as measured in Fig. 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 Fig. 4d (see Supplementary Videos 8–9), 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 50ms 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 50ms from –250ms to 500ms and then in steps of 100ms. Negative values indicate that the surround grating started to move after the dot changed its intensity. Additional time intervals were –500ms and –1s. The block experiments in Fig. 4f–j and Fig. S6 were performed

with the same frequency tuning stimuli as before (Fig. 4b). For the contrast tunings, the same stimuli as in Fig. 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.

Data evaluation. 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 interests (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/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[22]. 1D Gaussians were fit to horizontal and vertical cross-sections of spatial receptive fields to obtain precise RF coordinates. For lamina cells (Fig. S3), all reconstructed RFs were peak-aligned and analysed as previously[22]. For 1D projections of spatial RFs (Fig. S3f–j), an average of 1D projections of 2D RFs along 3600 evenly distributed projection angles between 0° and $360 s^{-1}$ was calculated. This enhanced the visibility of the center-surround structure but neglected possible anisotropies in the spatial structure of RFs[24]. For impulse responses (Fig. S3k–o) the temporal receptive field of the 9 center pixels was averaged; frequency responses (Fig. S3p–t) are the Fourier-transformed impulse responses. Deconvolution (Fig. S3u) was performed by dividing the frequency spectra with the frequency response of a 1st order low-pass filter with time-constant 350ms as a proxy for calcium indicator dynamics[22, 38].
Relative fluorescence changes ($\Delta F/F$) from raw calcium traces were obtained by adapting an automatic baseline detection algorithm[39]. Briefly, raw data were first smoothed with a Gaussian window (full-width at half maximum, $FWHM = 1s$). Then, minima within a 90s long sliding window were extracted and the resulting trace smoothed with a Gaussian window ($FWHM = 4min$). The result was used as a dynamic baseline $F0$ and $\Delta F/F$ values were computed as $\Delta F/F = (F - F0)/F0$.
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

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

where $s(t)$ denotes the signal and T the stimulation time. For experiments in Fig. 4d, we evaluated the peak response of the calcium signal.
Amplitudes were averaged over trials and normalized to the maximum, then averaged over cells and normalized to the maximum. For Fig. 4a–c,g,i and Fig. S6a,b, amplitudes were normalized to the response amplitude for the reference stimulus.

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[11] and then sampled individual signals from a rectangular grid with isotropic spacing of 4° (yielding 23 x 17 receptor signals per frame for data set A and 23 x 23 for data set B, as described above).
For the comparison in Fig. 1, we modelled 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 (Fig. 5) 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 $s.d. = 100^\circ s^{-1}$. The initial window phase followed a uniform distribution spanning 360°. To further augment the data set, 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 modelling experiments in Fig. 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 (Fig. S5b–d), the frequency experiment (Fig. S5f), and the background diameter experiment (Fig. S5g) were estimated from a single trial per parameter setting. For the background orientation experiment (Fig. S5e) and the step interval experiment (Fig. 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 (Fig. S5k) to account for the intrinsic stochasticity of the stimulus and to generate reliable model responses. Throughout Fig. 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 (Fig. 3v–x) resembles previous formulations in the literature[7, 40, 41]. 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 behaviour, 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 (Fig. 3x), this model was fit individually for each cell. Parameter fits to the average tuning curve per cell type are listed in Table S2. 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 $R^2_{DivisiveNorm}$ in Table S2). 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 S2).
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 Fig. 1 was based on a standard implementation of the Reichardt-type correlational motion detector[11]. Briefly, all receptor signals of the two-dimensional input grid (see above) were filtered with a first-order high-pass ($\tau = 150ms$). We then multiplied each local signal with the delayed horizontal neighbor (first-order low-pass, $\tau = 50ms$). 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 Fig. S1f, 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 (modelled after L3; Fig. S5b), a strongly normalized band-pass unit (modelled after Mi1; Fig. S5c), and a weakly normalized low-pass unit (modelled after Mi9; Fig. S5d). We hand-tuned parameters based on our and previous work[22] to qualitatively match response properties of the corresponding cell. Models were implemented as signal processing cascades (see Fig. 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 (Fig. 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 ($\tau = 80ms$) or serial low- ($\tau = 50ms$) and high-pass filters ($\tau = 150ms$) 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 (Fig. S5c,d).

To generate simulated T4 responses (Fig. S5i,j), 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 (Fig. S5k) as a spatial array of T4 and T5 cells covering the full field of view, analogously to the previously described two-quadrant detector[42]. 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 (Fig. S5l–n). All models in Fig. S5 were implemented using Python 3.6 and NumPy 1.15.

To quantify the robustness of velocity tunings for models and LPTCs (Fig. 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 x 17 x 500 or 23 x 23 x 500 (azimuth, elevation, time), depending on the data set, were processed in two independent convolutional channels. The convolutions were temporally causal and spatiotemporally separable. Each of the channels was composed of a 3 x 3 x 1 spatial filter (covering 3 simulated receptors in azimuth and elevation) followed by a temporal filter of shape 1 x 1x 30 (corresponding to 300ms 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")

$$f(x_i) = tanh(x_i/c)$$

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

$$f(x_i) = tanh(x_i/(c + P_i))$$

where c again determines the baseline sensitivity and P_i is the instantaneous output of a 11 x 11 x 1 spatial filter (centered on the location of x_i and operating on full-wave rectified output signals $|x_i|$; see Fig. 5a). 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, back-propagation, 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[43], 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 L2 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 hours on an NVIDIA Titan Xp GPU.

Statistics. 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. Normality of data distributions was assessed visually but not tested formally. Sample sizes 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.

Code availability. Code for analysis and modelling is available upon request to the authors.

Data availability. Data from behavior, physiology, and modelling are available upon request to the authors.

Methods References.

31. A. H. Brand and N. Perrimon, "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes," *Development*, vol. 118, pp. 401–415, June 1993.
32. T.-W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, and D. S. Kim, "Ultrasensitive fluorescent proteins for imaging neuronal activity," *Nature*, vol. 499, pp. 295–300, July 2013.
33. S. T. Sweeney, K. Broadie, J. Keane, H. Niemann, and C. J. O'Kane, "Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects," *Neuron*, vol. 14, pp. 341–351, Feb. 1995.
34. H. G. Meyer, A. Schwegmann, J. P. Lindemann, and M. Egelhaaf, "Panoramic high dynamic range images in diverse environments," 2014.
35. A. Bahl, G. Ammer, T. Schilling, and A. Borst, "Object tracking in motion-blind flies," *Nature neuroscience*, vol. 16, no. 6, p. 730, 2013.
36. E. Serbe, M. Meier, A. Leonhardt, and A. Borst, "Comprehensive characterization of the major presynaptic elements to the *Drosophila* OFF motion detector," *Neuron*, vol. 89, pp. 829–841, Feb. 2016.
37. T. A. Pologruto, B. L. Sabatini, and K. Svoboda, "ScanImage: flexible software for operating laser scanning microscopes," *Biomed. Eng. Online*, vol. 2, p. 13, May 2003.
38. F. G. Richter, S. Fendl, J. Haag, M. S. Drews, and A. Borst, "Glutamate signaling in the fly visual system," *iScience*, vol. 7, pp. 85–95, Sept. 2018.
39. H. Jia, N. L. Rochefort, X. Chen, and A. Konnerth, "In vivo two-photon imaging of sensory-evoked dendritic calcium signals in cortical neurons," *Nat. Protoc.*, vol. 6, pp. 28–35, Jan. 2011.
40. M. Carandini, D. J. Heeger, and J. Anthony Movshon, "Linearity and normalization in simple cells of the macaque primary visual cortex," *Journal of Neuroscience*, vol. 17, pp. 8621–8644, Nov. 1997.
41. T. C. B. Freeman, S. Durand, D. C. Kiper, and M. Carandini, "Suppression without inhibition in visual cortex," *Neuron*, vol. 35, pp. 759–771, Aug. 2002.
42. H. Eichner, M. Joesch, B. Schnell, D. F. Reiff, and A. Borst, "Internal structure of the fly elementary motion detector," *Neuron*, vol. 70, no. 6, pp. 1155–1164, 2011.
43. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

3 DISCUSSION

Drosophila melanogaster is an appealing model organism for the investigation of motion vision circuits. This is not only due to the fact that fruit flies show reliable and stereotypical motor reflexes in response to visual motion but also because the availability of a diverse neurogenetic toolset for circuit manipulation offers unparalleled opportunities to bring theoretical and computational approaches directly to the test bench for experimental verification.

For this thesis, I and my collaborators have applied some of these methods to several previously unsolved problems.

First, we systematically mapped the functional properties of all previously known input neurons to T4 and T5 cells using the same stimulus. Some of the inputs had been described before, but only subsets of them and investigated by means of different methods and stimuli. The medulla neurons Mi4 and Mi9 had not been functionally characterized before I started working on my doctoral projects. Using computer simulations, we could show that Mi4 and Mi9 would be ideally suited to constitute the two delayed arms of a three-arm hybrid motion detector scheme due to their distinct temporal filtering characteristics.

Furthermore, we looked at glutamate signaling properties in the fly motion vision circuitry. We confirmed the glutamatergic identity of Mi9 neurons. By using a glutamate indicator which has markedly faster kinetics than common calcium indicators we could measure temporal processing properties of this cell type with significantly higher precision. The results were in line with our predictions from the first study which were based only on post-hoc deconvolution of calcium responses with a putative calcium indicator kernel.

In another study, we showed that a class of visual projections neurons, LC10, computes relative motion in order to relay positional information about conspecific flies and possible mating partners to courtship circuits in the male brain. This is a compelling example of how local motion cues can be exploited for extraction of more complex visual features.

Finally, we showed that input neurons to T4 and T5 adapt their dynamic range to the prevailing contrast level of the stimulus. Implementation of a contrast adaptation stage renders models of the fly motion vision circuitry robust to the natural contrast fluctuations pervasive in natural images. The mechanism resembles divisive contrast normalization, a neural computation which has been described abundantly in the vertebrate visual system. This is the first demonstration of contrast normalization in the invertebrate visual system and hence provides another example for evolutionary convergence in neural networks. Our findings represent an important step towards understanding robust motion processing of realistic inputs.

Most of the calcium imaging and electrophysiology experiments described in this thesis have been conducted in tethered fruit flies placed in a cylindrical projector-based stimulation arena for display of visual stimuli. The design of this stimulation arena, in a way that it can be replicated for standardization of visual stimuli across experimental set-ups, was a major part of my doctoral work. Additionally, I programmed a suitable software front-end for control and design of visual stimuli as well as for synchronization of the experimental data with the stimulus timing and subsequent data analysis.

Here, I will discuss the scientific implications of the work that has been presented in this thesis in light of more recent research published after I started my doctoral work. Moreover, this chapter will also provide a short summary about function, advantages and limitations of the new visual stimulation system.

3.1 THE EMERGENCE OF DIRECTION SELECTIVITY

3.1.1 Non-linear interactions

Standard models used to explain the emergence of direction selectivity have been based on either an enhancing or a suppressive non-linear interaction between two spatially offset input arms with distinct temporal processing properties (see Figure 8). As elaborated above (see Section 1.2.2), calcium imaging experiments in T4 led to the proposal of the hybrid detector which combines both types of non-linear interactions into a three-arm model of motion detection (see Figure 9) (Haag et al., 2016; Leong et al., 2016). Recently published electrophysiological recordings of T4 cells in response to apparent motion stimuli, however, did not find an enhancing non-linearity in T4 voltage signals and thus argued for pure inhibitory interactions to account for direction selectivity (Gruntman et al., 2018). The difference between those studies could, however, be attributed to the size of the stimulus used. In a follow-up study, Haag et al. (2017) titrated the size-dependence of preferred direction enhancement in T4 and found significant supra-linear interactions only for stimuli larger than about 4-5°. Since Gruntman et al. (2018) used only ~2° wide bars, this might explain why they did not observe preferred direction enhancement in T4. Another possibility is that the non-linear enhancement clearly seen in calcium imaging experiments originates from non-linear transformation of voltage to calcium signals. Since calcium concentration is directly related to neurotransmitter release in a synapse, non-linearities in this transformation might indeed be of physiological relevance for the effective output signal of T4 cells.

In the OFF-pathway, similar experiments have now been conducted and have confirmed the co-occurrence of preferred direction (PD) enhancement and null direction (ND) suppression in calcium signals also for T5 cells (Haag et al., 2017). Another study used a voltage indicator in order to measure T5 responses to stationary and moving sinusoidal gratings, a stimulus protocol which allows for linear systems analysis in a similar way as with apparent motion stimuli (Wienecke et al., 2018). This study found voltage responses in T5 to be completely described by a linear model based

on a spatiotemporally tilted linear receptive field and attributed all non-linearities to the voltage-to-calcium transformation. Gruntman et al. (2018) pointed out that bilobed and tilted spatiotemporal receptive fields need not necessarily result from a hybrid mechanism, but can also result from asymmetric inhibition of an excitatory subfield alone. On the other hand, these options are not mutually exclusive: lateral excitatory inputs could still add to the excitatory subfield in order to shape the space-time orientation of the receptive field and thus increase direction selectivity. However, it is not yet clear if such linear-nonlinear model would be able to account quantitatively for all observed T5 response properties at the same time. Electrophysiological recordings in T4 have clearly argued for intrinsic non-linearities already at the level of membrane voltage at least regarding ND suppression. Similar experiments need to be performed in T5 in order to address the uncertainties regarding the nature of non-linear interactions in these cells.

Ultimately, in order to fully understand the emergence of direction selectivity in T4 and T5, one has to consider also the biophysical and anatomical constraints that are put on our models of motion detection. Does the anatomical arrangement of input neurons match their temporal filtering properties according to models of motion detection? Which neurotransmitters are released by the input elements? How does activation or silencing of these inputs affect T4 and T5 neurons? Only these questions can eventually pinpoint the exact mechanism of how direction selectivity emerges on T4 and T5 dendrites.

3.1.2 Spatiotemporal receptive fields of the input cells

Asymmetric temporal filtering of spatially offset input signals is a constituting element of motion detector models. While it has been proposed that temporal delays could be implemented via slow receptors on T5 dendrites (Shinomiya et al., 2014), most studies on fly motion detection have argued for an implementation through intrinsically different temporal filtering properties of the input neurons to T4 and T5 (Behnia et al., 2014; Serbe et al., 2016; Strother et al., 2017). The two possibilities are not mutually exclusive, but further insights into the synaptic mechanisms between presynaptic partners and T4 and T5 are necessary. In contrast, a lot is known about the functional properties of the input elements.

Before I started my doctoral work, functional characterization had already been performed for subsets of input neurons, but using different methods and stimuli. In the ON-pathway, spatial and temporal filtering properties of Mi1 and Tm3 cells had been measured using electrophysiology and white noise stimuli (Behnia et al., 2014) as well as using voltage imaging and full field flicker stimuli (Yang et al., 2016). However, Mi4 and Mi9 neurons had not been functionally described before. In the T5-pathway, Meier et al. (2014) and Serbe et al. (2016) had characterized Tm1, Tm2, Tm4 and Tm9 in calcium imaging experiments using flickering local bars of different sizes and durations. While they found all of these cell types to possess local receptive fields with inhibitory surrounds, Fisher et al. (2015a) described

Tm9 as a wide-field neuron using stochastic stimuli. Electrophysiological studies had been conducted for Tm1 and Tm2 using white noise (Behnia et al., 2014) as well as voltage imaging experiments using large field flicker (Yang et al., 2016).

Although most of these studies are in qualitative agreement, it is hard to compare results quantitatively due to the specific drawbacks of each method. Limited temporal resolution and intrinsic non-linearities of calcium and voltage indicators inevitably lead to distortions in imaging experiments. On the other hand, electrophysiology might not reflect the underlying axonal responses because of the large distance between axons and cell bodies in *Drosophila*. The use of different stimuli is another source of uncertainty: Non-linear cell properties will be reflected in step responses to flicker stimuli but not in receptive field reconstructions from stochastic stimuli since reverse correlation extracts only linear response components. Additionally, neural adaptation processes such as adjustment of luminance sensitivity or contrast gain control might lead to differentially adapted cells for different stimuli. Overall, in order to build realistic circuits models involving all cell types, there was the need for a dataset comprising a functional characterization of all input neurons assessed on equal grounds.

In Manuscript 1, we followed a systematic approach for mapping the spatiotemporal receptive fields of all previously known input neurons to T4 (Mi1, Tm3, Mi4, Mi9) and T5 (Tm1, Tm2, Tm4, Tm9) using calcium imaging and white noise stimuli. We find local receptive fields for all neurons, confirming earlier results (Behnia et al., 2014; Serbe et al., 2016) but in contrast to Fisher et al. (2015a) regarding Tm9. Interestingly, all neurons except of Tm3 show pronounced center-surround antagonism in their spatial receptive fields. With respect to their temporal processing properties, they seem to subdivide in two main classes: While Mi1, Tm3, Tm1, Tm2 and Tm4 have transient band-pass filter characteristics, Mi4, Mi9 and Tm9 exhibit low-pass filter properties. Differences among the elements of one class, e.g. between band-pass filter neurons Mi1 and Tm3, are minute in comparison with differences between these two distinctive main classes.

Intriguingly, low-pass elements seem to possess a stronger spatial surround, virtually cancelling out the center completely. Similar results have been obtained in more recent studies by measuring step responses to flashing local discs or bars (Strother et al., 2017; Salazar-Gatzimas et al., 2018). It will be interesting to investigate within the framework of predictive coding and redundancy reduction whether for these cell types the lack of temporal antagonism due to their low-pass characteristics is related to an enhancement of spatial antagonism.

The polarities of the receptive fields match the assignment of cell types to either ON- or OFF-pathway, with the exception of Mi9 which possesses an OFF-center. We found Mi9 axon terminals to co-localize with the vesicular glutamate transporter VGlut indicating its glutamatergic phenotype (see Manuscript 3). Sign-inversion through glutamate-gated ion channels could therefore render Mi9 effectively an ON-cell. The polarity of Mi9 could then be explained by reciprocal inhibitory connections between Mi4 and Mi9 (Takemura et al., 2013). Indeed, a recent connectomics study revealed the GABAergic identity of Mi4 as well as confirmed Mi9 to express VGlut

and its reciprocal connectivity with Mi4 (Takemura et al., 2017). Although the polarity of Mi9 can be explained in light of these findings, the specific functional advantage of such sign-inversion remains intriguing.

3.1.3 Anatomical arrangement of the input elements

Do the functional characteristics of the input elements map onto models of motion detection with respect to their anatomical arrangement along T4 or T5 dendrites?

Under the premise of a three-arm hybrid detector of motion detection, we tested the direction selectivity of different anatomical arrangements of input elements for both ON- and OFF-pathway (see Manuscript 1). Our computer models were agnostic to the sign of the synapses because neurotransmitter information was not yet available at that time. For T4, our simulations predicted that primarily configurations with low-pass filters, e.g. either Mi4 or Mi9 on the lateral arms of a three-arm detector would result in a high degree of direction selectivity. For T5, we found that placing the low-pass filter Tm9 on the suppressive side and Tm2, which is mildly faster than Tm1 and Tm4, in the center resulted in a high direction selectivity index.

Overall, we note that maximizing the difference of temporal filtering properties between neighboring input channels leads to high direction selectivity of simulated motion detectors. This is in stark contrast to previous studies which argued for two arm models of motion detection despite only small temporal delays between the input channels (Takemura et al., 2013; Behnia et al., 2014). None of these studies, however, addressed the requirement of high directional selectivity already at the level of the half-detector before the subtraction stage (see Figure 8). Measuring the model output after subtraction of two mirror symmetric half-detectors always yields highly direction selective responses regardless of the performance of the half-detectors (Haag et al., 2016).

Are our predictions supported by the actual connectivity in the fly medulla? Previous electron microscopic reconstructions comprised too small volumes to reliably map cross-column connections or trace the identity of neurons through the internal optic chiasma (Takemura et al., 2013; Shinomiya et al., 2014). In order to overcome these issues, recent electron microscopy studies have reconstructed larger volumes and hence provided a more complete connectivity matrix of the motion vision pathway in *Drosophila* (Takemura et al., 2017; Shinomiya et al., 2019).

These efforts have revealed that input synapses of different cell types are highly segregated along the T4 dendrite and provide spatially offset input signals from neighboring columns. While Mi9 synapses cluster on the tips of the dendrite, Mi1 synapses accumulate in the center region and Mi4 synapses are mostly found towards the base of the dendrite (see Figure 14 a). This is in line with our theoretical considerations that Mi9 or Mi4 should be placed laterally in a three-arm detector. Additionally, the reconstructions have shown input synapses from C3 and CT1 cells onto the base of the T4 dendrite. Intriguingly, they also found synaptic connections from neighboring T4 cells of the same subtype onto the tips of the dendrite.

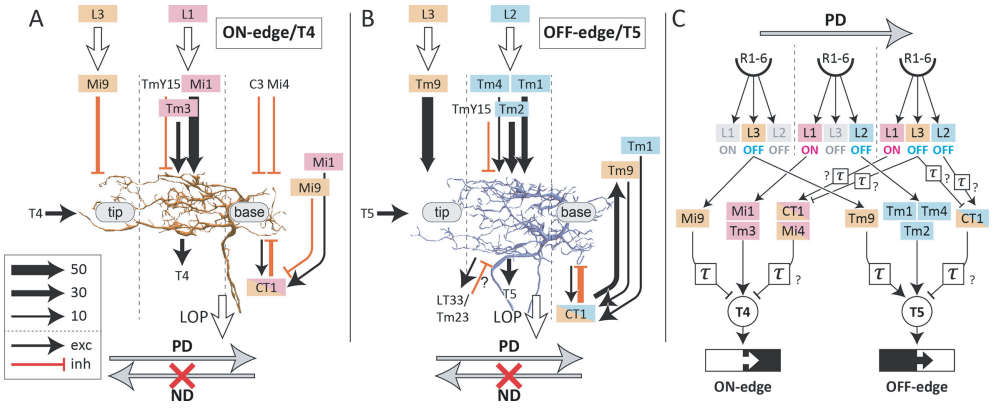


Figure 14: Spatial arrangement of synaptic connections on T4 and T5 dendrites | **a**, Schematic of the spatial distribution of input synapses along the T4 dendrite. The dendrite is partitioned into three compartments (dashed lines). **b**, Same as in a, for the T5 dendrite. **c**, Schematic of a possible implementation of a three-arm hybrid detector for both pathways, as inspired by the input synapse spatial distributions found on T4 and T5 dendrites. Graphics taken from Shinomiya et al. (2019) / CC BY.

For T5 cells, clustering of Tm9 synapses on the tips of T5 dendrites has been confirmed, as previously described (Shinomiya et al., 2014, 2019). However, Tm1 synapses were shown to aggregate together with Tm2 and Tm4 synapses in the central compartment of the dendrite. Additionally, and in analogy to T4 cells, Shinomiya et al. (2019) found reciprocal CT1 synapses on the base of the T5 dendrite and input from syndirectional T5 cells on the tips of the dendrite (see Figure 14 b). This arrangement is not in agreement with our predicted configuration from computer simulations because here Tm9 was placed on the position of the suppressive arm corresponding to the base of the dendrite. However, the finding of Tm9 synapses on the tips of the dendrite conforms with our notion that temporally differentially tuned inputs should be placed on separate arms of the detector. CT1 was not yet known to be a prominent synaptic input partner cell to T5 at the time of our first study.

Overall, the anatomical arrangement of input neurons along T4 and T5 dendrites matches their expected distribution in a three-arm model of motion detection (see Figure 14c). However, a functional description of CT1 (see next section) and C3 cells is necessary in order to test if these cell types fit into the scheme as well. In addition, the model postulates quite specifically the sign of the respective synaptic connections. The next section will provide an overview over recent insights into synaptic and molecular aspects of this circuit.

3.1.4 Synaptic mechanisms of motion detection

ON-pathway

As indicated above, immunolabeling and transcriptomics experiments suggest that Mi9 uses glutamate as a neurotransmitter (see Manuscript 3) (Takemura et al., 2017; Davis et al., 2018). Sign-inversion at the Mi9-T4

synapse could therefore switch the effective polarity of Mi9 postsynaptic signals. Such sign-inverting synapse would require the presence of a glutamate gated chloride channel in T4. Indeed, there is increasing evidence that T4 neurons express *GluClα* channels (S. Fendl, personal communication) (Pankova and Borst, 2016; Davis et al., 2018). *GluClα* has been found to mediate glutamatergic inhibition in *Drosophila* and other invertebrates (Cleland, 1996; Liu and Wilson, 2013; Mauss et al., 2014, 2015). An implementation of preferred direction enhancement through release from inhibition instead of simple excitation has been suggested to lead to non-linear signal amplification of subsequent excitatory signals by increasing the input resistance in a passive membrane model of a three arm detector (Borst, 2018).

In the central column, T4 is predominantly contacted by Mi1 and Tm3 cells, which have been shown to provide cholinergic excitatory input (Pankova and Borst, 2017; Takemura et al., 2017). This fits well into a three-arm hybrid model of motion detection.

Finally, immunolabeling results suggest that ND suppression in T4 could be implemented through GABAergic inhibition from Mi4 synapses at the base of the dendrite (Takemura et al., 2017). Additional suppression might come from GABAergic CT1 neurons (Takemura et al., 2017). CT1 is a giant tangential cell that sends it neurites to every column of the lobula and of the medulla. Only recently, it has been shown that this cell type exhibits such a high degree of compartmentalization that each neurite terminal can be seen as an independent computational compartment. Individual terminals possess distinct and local receptive fields, even though all of them belong to the same CT1 cell (Meier and Borst, 2019). Intriguingly, CT1 receptive fields in the medulla are ON-selective while they have OFF-polarity in the lobula. This renders CT1 a good candidate for null direction suppression via synaptic inhibition in both T4 and T5 cells. In addition, C3 cells provide GABAergic inhibition on the same side of the T4 dendrite (Kolodziejczyk et al., 2008; Takemura et al., 2017). A functional characterization of this cell type is still missing.

The anatomical arrangement of input cells corresponds conceptually well to a three-arm model of motion detection. Yet, direct proof for the validity of this model would have to come from activation or silencing experiments. Blocking synaptic output of Mi1 or Tm3 cells using genetically targeted expression of $shibire^{ts1}$ leads to disruption of drifting grating and edge responses in T4 cells and of optomotor behaviour (Strother et al., 2017). This resembles previous findings of Ammer et al. (2015) and matches expectations since blocking the central arm of the detector should heavily impair its function. However, silencing Mi4 or Mi9 did not diminish direction selective responses in T4, but had modulatory effects on temporal tunings (Strother et al., 2017). Blocking T4-T4 synaptic connections did not have discernible effects on T4 direction selectivity (Haag et al., 2016). Besides ineffective silencing tools, a possible explanation for this lack of a blocking effect for these cell types might be the fact that excitation as well as inhibition are implemented redundantly by more than one cell type on either side of the T4 dendrite. To address this possibility, future experiments should also test for combinatorial blocks of different cell types. Another problem is that blocking only one side of the detector is not expected to lead

to complete abolishment of T4 responses but only to partially decreased direction selectivity. The use of apparent motion stimuli instead of gratings or edges in blocking experiments might therefore be more instructive here. Recording responses to apparent motion in T4 while blocking specific input cell types could distinguish between effects on only supra-linear or sub-linear response components in T4.

Another experimental approach is to activate combinations of input cell types in order to pinpoint pairs of neurons which interact non-linearly on the T4 dendrite. Strother et al. (2017) optogenetically activated combinations of either Mi1 or Tm3 together with Mi1, Tm3, Mi4 or Mi9 cells using the red-shifted channelrhodopsin *Chrimson*. They observed substantial supra-linear integration of Mi1 and Tm3 signals in T4 calcium responses. Since a motion detector based on only these two cell types has already been ruled out by anatomical analysis, the observed non-linear interactions are likely due to other intracellular mechanisms or to the voltage-to-calcium transformation. Activating either Mi4 or Mi9 together with Mi1 or Tm3 did not lead to supra-linear responses which is expected if Mi4 and Mi9 are inhibitory. In order to test for supra-linear interactions between Mi9 and Mi1 or Tm3, one would have to hyperpolarize Mi9 but depolarize the other cell type which might be possible by appropriate choice of spectrally separable optogenetic agents.

OFF-pathway

In the OFF-pathway, central excitation seems to be provided by Tm1, Tm2 and Tm4 which all have been characterized as cholinergic neurons (see Figure 14b,c) (Takemura et al., 2011; Shinomiya et al., 2014). PD enhancement might be implemented through cholinergic Tm9 cells which provide input from the neighboring column (Shinomiya et al., 2014). Thus, sign-inversion as in the ON-pathway is not necessary here. Supra-linear interactions between Tm9 and the central inputs could then be due to voltage-gated ion channels. Alternatively, central inputs could be modulated by a ligand-gated ionotropic receptor under control of a second messenger cascade downstream of a G-coupled receptor between Tm9 and T5. Instead of a G-coupled receptor, another possibility is that ligand-gated calcium channels could regulate the sensitivity of receptors for central input signals.

Single cell blocks of Tm1, Tm2 or Tm4 did not have strong deteriorating effects on LPTC responses to OFF-edges (Serbe et al., 2016). This can be readily explained by their redundant function in the center of the motion detector. Silencing of Tm9 partially reduced LPTC responses and strongly diminished T5 calcium responses arguing for an important role of Tm9 in OFF-edge motion detection (Fisher et al., 2015a; Serbe et al., 2016). This is accompanied by combinatorial block results which show that double cell blocks involving only Tm1, Tm2 or Tm4 cells only moderately impairs OFF-edge LPTC responses, while combinations also involving Tm9 have the strongest effect (Serbe et al., 2016). Similar experiments need to be performed using T5 calcium or voltage signals as a read-out.

With synapses located at the base of the T5-dendrite, CT1 could provide inhibition for ND suppression which is in line with the OFF-polarity of CT1 terminals in the lobula (Meier and Borst, 2019). There is no other inhibitory

cell type with a similar spatial synapse distribution along T5 dendrites. Therefore, it is expected that blocking CT1 should effectively abolish ND suppression in T5.

Optogenetic activation experiments have not yet been conducted in the OFF-pathway.

Outlook

While the anatomical arrangement of input cells and their neurotransmitter profiles strongly speak in favor of a three-arm hybrid mechanism in both pathways, it is not yet clear which specific biophysical mechanisms lead to non-linear interactions between the input channels. It has been shown that ND suppression can result from inhibition in a passive conductance-based model of the T4 dendrite (Gruntman et al., 2018). Importantly, the morphology of the dendrite does not seem to play a substantial role other than collecting inputs from several columns in this model. Borst (2018) pointed out that ND suppression could be implemented either through shunting inhibition or via direct inhibition followed by a threshold function such as that of a calcium channel. The exact nature of the interaction depends on the magnitude of the inhibitory conductance change in relation to the leak conductance of the cell. In the same study, it has been suggested that PD enhancement on the other side of the T4 dendrite could result from supra-linear interactions following a release from inhibition upon deactivation of Mi9 by an ON-stimulus. This requires that Mi9 is constantly inhibiting T4 in the absence of an ON-stimulus. Evidence for that could be obtained from glutamate imaging or electrophysiological experiments. In the OFF-pathway, Wienecke et al. (2018) argue that all non-linearities are due to the voltage-to-calcium transformation. While this is entirely possible, the study did not build an explicit model for this transformation from voltage recordings to calcium data. Ultimately, a complete electrophysiological characterization of T4 and T5 cells comprising membrane voltage dynamics but also changes in the excitatory and inhibitory conductances upon presentation of a diverse set of stimuli will be necessary in order to fully address these questions and to gain deeper insight into the biophysical mechanisms of motion detection.

3.1.5 Behavioral state modulation

Flies adapt their visual processing properties depending on their current behavioral state. For example, response gain of HS cells is increased in walking flies, and their temporal frequency tuning is shifted towards higher frequencies when compared to stationary flies (Chiappe et al., 2010). This adjustment of tuning curves could serve to better match the expected stimulus statistics of visual input signals during active locomotion. Similarly, VS cells in *Drosophila* exhibit higher gain during tethered flight (Maimon et al., 2010). Furthermore, it has been shown that active flight boosts the responses of wide-field horizontal motion selective H1 cells at high velocities in tethered flying *Lucilia* (Jung et al., 2011). This effect can be reproduced by a reduction of the time constant in models of motion detection depending

on the behavioral state. Pharmacologically, the effect can be mimicked by applying the octopamine agonist chlordimeform (CDM) to the brain of the fly. In a similar way, a boost of VS cell responses can be triggered by activation of octopaminergic neurons using the temperature sensitive cation channel *dTrpA1* (Suver et al., 2012). Octopaminergic neurons have been suggested to be involved in diverse modulatory functions across the whole nervous system of the fly (Sinakevitch and Strausfeld, 2006; Busch et al., 2009). In the optic lobe, specifically, it has been shown that octopaminergic neurons are active during flight (Suver et al., 2012). It is therefore believed that state-dependent modulation of motion detection circuits in the fly is mediated by the octopaminergic system.

In Manuscript 1, we showed that behavioral state modulation acts already on early visual neurons in *Drosophila*. First, we found that T4 and T5 neurons shift their velocity tunings towards higher velocities upon pharmacological intervention using CDM. Second, we scanned the spatiotemporal receptive fields of T4/T5-presynaptic medulla neurons after application of CDM. While spatial receptive fields remained identical, temporal processing properties of all cell types significantly increased in speed. This effect was particularly pronounced for the band-pass filter elements Mi1, Tm3, Tm1, Tm2 and Tm4, but also moderately affected low-pass filter cells Mi4, Mi9 and Tm9. A recent study confirmed these findings in the ON-pathway (Strother et al., 2018). This study could replicate the effect not only by pharmacological intervention, but also relate it to the state of locomotion of the fly. Intriguingly, behavioral state modulation is exceedingly strong in Mi4 cells. Using optogenetics in a presynaptic lamina neuron, L5, Strother et al. (2018) could show that application of CDM increases the excitability of Mi4 cells substantially, thereby demonstrating a possible physiological mechanism for behavioral state modulation.

In total, these findings corroborate the original suggestion by Jung et al. (2011) that velocity tunings of motion sensitive neurons in the fly can be adjusted by means of adaptive temporal processing properties in visual neurons presynaptic of T4 and T5 cells. The modulatory control over these neurons is exercised through the octopaminergic system in dependence on the behavioral state of the animal. Of course, it cannot be excluded that state-dependent modulation acts on different levels of the motion vision circuit at the same time. It will be interesting to dissect the exact mechanisms of how octopaminergic signaling can modulate temporal processing features of visual interneurons in more detail.

3.2 CONTRAST NORMALIZATION

3.2.1 Adaptation and non-linear preprocessing

A large part of my doctoral work was dedicated to the detailed description of *linear* spatiotemporal filter properties of input neurons to T4 and T5 cells (see Manuscript 1 and Manuscript 3). I then used these insights to test possible circuit implementations of motion detection, thereby basing my computer simulations on purely linear input filters. This approach

was viable, because in this work we investigated mechanisms of motion detection exclusively by use of artificial stimuli with full contrast. However, models of motion detection with linear input channels are known to have poor performance at estimating motion from naturalistic signals (Dror et al., 2001; Brinkworth and O'Carroll, 2009; Fitzgerald and Clark, 2015; Leonhardt et al., 2016). This is because correlation-type motion detectors are prone to confuse image contrast with scene velocity. For example, in a model like the Hassenstein-Reichhardt detector, the output depends on the temporal overlap of the two asymmetrically filtered input signals as well as on their contrast in a quadratic way due to the multiplicative interaction (see Section 1.1.4). Natural images are highly variable and exhibit strong local contrast fluctuations (Geisler, 2008). As a result, correlation-based motion detectors without appropriate signal preprocessing will reflect this variability.

On a purely computational level, this issue can be partially addressed by introducing non-linear preprocessing stages into the models. Compressive non-linearities and response saturation, local motion adaptation, exploitation of ON/OFF-asymmetries or combinations of these have successfully been applied in models of motion detection to improve performance with naturalistic input signals (Dror et al., 2001; Shoemaker et al., 2005; Brinkworth and O'Carroll, 2009; Fitzgerald and Clark, 2015; Leonhardt et al., 2016). Saturation in the input lines can also account for the robustness of HS cell responses given synthetic stimuli superimposed with random pixel noise (Suzuki et al., 2015). However, these approaches study the problem only from a theoretical perspective without pinpointing corresponding biological mechanisms, with the exception of the study by Leonhardt et al. (2016) where ON/OFF-asymmetries in temporal signal processing have successfully been verified in the fly motion vision pathway.

In the biological system, various mechanisms of adaptation to natural scene statistics have been suggested which help alleviate this problem. Gain control in fly photoreceptors is essential in order to map the large range of intensities in natural environments onto the limited dynamic range of neurons (van Hateren, 1997). Center-surround receptive fields and temporal inhibition in lamina monopolar cells reduce redundancy and eliminate neighboring pixel correlations inherent to natural scenes (Srinivasan et al., 1982). Finally, as mentioned above, ON- and OFF-channels in fly motion vision are differentially tuned in order to exploit natural ON/OFF-asymmetries (Leonhardt et al., 2016). However, none of these mechanisms addresses specifically the strong contrast fluctuations across and within natural scenes, to which correlation-type motion detectors are particularly vulnerable.

3.2.2 Gain control in the fly

In Manuscript 4, we described a mechanism for contrast gain control which acts in the medulla of the fly visual system. This mechanism suppresses neuronal responses to local contrast divisively in dependence of the average scene contrast, estimated over a spatially extended area in

visual space. Thus, the mechanism effectively normalizes responses to the prevailing average contrast of the stimulus. Contrast normalization is found only in a subset of neurons in the medulla (Mi1, Tm1, Tm2, Tm3, Tm4) and is inherited by T4 and T5 direction selective cells. Intriguingly, this subset comprises all transient cells in the medulla but excludes tonic low-pass filter elements Mi4, Mi9 and Tm9. Lamina neurons L1–5 are not subject to contrast normalization. Therefore, contrast normalization emerges in the medulla. The effect of normalization is reflected in optomotor behavior of tethered flies and thus it has behavioral relevance.

Using a task-driven modelling approach exploiting fully differentiable convolutional neural networks, we compared different strategies of signal preprocessing. First and in agreement with previous studies, we find that any kind of static signal compression already improves the performance of motion detector models with naturalistic input stimuli over linear transmission of input signals. However, dynamic contrast gain adaptation, as we find in the fly, is superior to a static non-linearity in the input lines and closes the performance gap between the fly and the model. Our work represents therefore an important step towards understanding robust sensory processing of natural input signals with respect to motion vision in the fly.

The described mechanism closely resembles divisive normalization, as it is known from e.g. cortical processing, the LGN and also the fly olfactory system (Olsen et al., 2010; Carandini and Heeger, 2011). To my knowledge, this is the first demonstration of such mechanism in the fly visual system. Moreover, we provide a comprehensive characterization of neuronal contrast tunings throughout the whole motion vision circuitry. Our findings will therefore have an impact on research regarding the mechanism of direction selectivity in the fly. For example, since early studies it has been argued that the contrast dependency of direction-selective cells should be determined by the mechanism of motion detection in the fly (Poggio and Reichardt, 1976). Our study shows that the contrast tunings of T4 and T5 cells are mainly given by their input elements and that they inherit gain control properties from their inputs.

Mechanisms of gain control in the fly visual system have been described earlier. For example, H1 neurons in *Lucilia* adapt strongly over time to moving stimuli (Maddess et al., 1985). Interestingly, electrophysiological recordings from HS cells in the dronefly revealed later that this type of motion adaptation is mediated by a change in the contrast sensitivity of the system (Harris et al., 2000). This contrast gain control depends on motion, but is insensitive to the direction of the stimulus. In that, it resembles our characterization of the normalization signal in the fly medulla, which is independent of the direction of background motion. However, we describe a gain control signal that is elicited by dynamic contrast and not necessarily by motion. Moreover, the adaptation described by Harris et al. (2000) acts in the temporal domain and is induced through prolonged presentation of an adapting stimulus over several seconds. In contrast, normalization in the fly medulla depends on a spatial assessment of average contrast in the vicinity of the receptive field and acts within tens of milliseconds or faster.

A spatial gain control mechanism in the fly has been suggested after measurements of optomotor behavior in *Calliphora*, which showed that the magnitude of the optomotor response is invariant to the size of an oscillating figure (Reichardt et al., 1983). Such size-invariance can be explained by a divisive gain control mechanism that normalizes the response to the average motion content of the scene. Similarly, size-invariant tunings for windowed motion stimuli have also been reported for HS and VS cells in large flies (Hausen, 1982; Egelhaaf, 1985; Haag et al., 1992). To explain these phenomena, originally, Reichardt et al. (1983) proposed a binocular pool cell mechanism that averages the activity of all elementary motion detectors and in turn provides shunting inhibition to each of them. This is conceptually close to our suggested mechanism of contrast normalization, but differs in that it only accounts for contrast gain control of motion-selective and downstream neurons. Gain control of this type, in these early studies also often referred to as *motion adaptation*, was later attributed to non-linear dendritic integration mechanisms in LPTCs (Borst et al., 1995; Weber et al., 2010). It is important to note that such models predict gain control to depend primarily on visual motion, even if the gain control signal could be independent of the specific direction of motion because the pool mechanism isotropically averages direction selective units with all preferred directions. This is in contrast to our findings which show that the optomotor response in *Drosophila* is subject to gain control also for spatiotemporally uncorrelated contrast fluctuations.

Borst et al. (2005) pointed out that HR detectors exhibit intrinsic gain control properties emerging from the fundamental non-linearity of the system. This leads to an adaptive velocity gain reduction of the system, which depends on the amplitude of velocity fluctuations in the input stimulus. Again, this form of gain control is dependent on intrinsic spatiotemporal correlations of the stimulus and is not controlled by the average contrast of the stimulus.

A gain control mechanism based on spatial assessment of average scene contrast has not been described in the fly before our study. It will be interesting to reassess previous findings of gain control in fly motion processing and investigate more in detail to which degree contrast normalization already in the medulla could provide an explanation for these phenomena. For instance, non-linear spatial integration of motion cues, as observed in LPTCs, could also results from a network model with normalized medulla neurons. However, this does not exclude that non-linear dendritic integration, as proposed by Borst et al. (1995), contributes to the phenomenon as well, if excitatory and inhibitory conductances are comparable in magnitude to the leak conductance of the cell. To quantify the relative contribution of each mechanism, precise measurements of the synaptic and leak conductances of LPTCs in *Drosophila* are necessary.

3.2.3 Comparison with vertebrate vision

Our work gives an example for why it is of great advantage to preprocess signals using response normalization prior to extraction of higher-order

features, such as the direction of motion in T4 and T5 cells in our case. This mechanism is known from vertebrate visual processing, and contrast normalization in the fly shares many attributes with normalization in the cortex or LGN, but there are also significant points in which they differ.

Broadly speaking, in vertebrate vision there is a distinction between *surround suppression* and *cross-orientation suppression*. Surround suppression is a form of divisive modulation in simple or complex cells in V1, which can be elicited by stimuli outside of the classical receptive field (Cavanaugh et al., 2002a,b). It is maximal when the suppressive stimulus has the same orientation as the preferred orientation of the cell (Cavanaugh et al., 2002b; Carandini and Heeger, 2011). In contrast to that, cross-orientation suppression mainly originates from within the classical receptive field but is largely independent of the orientation of the stimulus (DeAngelis et al., 1992; Carandini et al., 1997; Carandini and Heeger, 2011). Similarly, neurons in the LGN have been shown to be modulated by a suppressive signal that is independent of orientation and originates from a summation field mainly within the classical receptive field (Bonin et al., 2005).

As in the fly, divisive suppression in the LGN is thought to originate from a pool mechanism that computes the standard deviation of the input signals as an estimate for average contrast. In fact, the suppressive signal in the LGN has been shown to be rigorously coupled to the root-mean-square contrast of the stimulus (Bonin et al., 2006). As in the fly, the suppressive signal in the LGN is direction independent (Bonin et al., 2005). In both systems, temporal frequency tunings are band-pass. In the LGN, the suppressive signal is broadly tuned to temporal frequency but exhibits a peak at around 10-20 Hz. In the fly, peak suppression lies at around 2 Hz which is substantially lower than in the LGN. Furthermore, divisive suppression in the fly is not effective for static stimuli. A major difference between the two systems is that contrast normalization in the fly medulla originates from spatial integration of an area significantly larger than the classical receptive field of a typical medulla neuron, which is in contrast to the LGN where it is constrained to the area of the classical receptive field of an LGN neuron.

What is the functional significance of response normalization in vertebrate visual processing? In early studies, divisive contrast gain control has been linked to dynamic range compression, serving to map the wide range of contrasts encountered in natural environments onto the limited coding range of neurons (Heeger, 1992). In early visual processing, such as in the retina, the main rationale for contrast gain adaptation is dynamic sensitivity adjustment (Shapley and Victor, 1978; Baccus and Meister, 2002). As pointed out in the introduction, contrast normalization has also been suggested as a way to reduce statistical redundancy of naturalistic input stimuli and to increase the efficiency of their neural representations (Schwartz and Simoncelli, 2001). In the cortex, it has been shown that normalization can perform a winner-take-all computation between concurrent stimuli in a population of neurons (Busse et al., 2009). Overall, there are various proposals for the role of normalization in vertebrate visual processing, most of which are related to general coding principles (Carandini and Heeger, 2011), and it is also possible that the mechanism fulfills several functions at the same time.

In our study in the *Drosophila* visual system, we were able to relate the function of contrast normalization directly to a behavioral task which is critical for the survival of the animal: the precise and reliable estimation of velocities from moving natural images. Contrast normalization as a strategy for signal preprocessing in the medulla allowed our model of the fly motion vision circuitry to achieve fly-like performance at this task. Hence, contrast normalization plays a critical role to render motion processing pathways in the fly invariant to the statistical fluctuations of contrast in natural environments.

There is not much known about the neural mechanisms that implement contrast gain control in the cortex. Measurements of the precise onset and offset dynamics of suppression in V1 revealed that cross-orientation suppression is significantly faster than surround suppression (Bair et al., 2003; Smith et al., 2006). This might indicate that cross-orientation suppression is mediated via feedforward connections. Freeman et al. (2002) suggested synaptic depression at the thalamocortical synapse as a candidate mechanism for cross-orientation suppression. However, a later study argued that the suppressive signal is too rapid for synaptic depression and might instead be due to nonlinear properties of LGN neurons or other upstream processes (Li et al., 2006). Similarly, it was proposed that also LGN divisive suppression might be inherited by upstream retinal properties and only be further enhanced in the thalamus (Bonin et al., 2005). However, the relative contributions of each of the processing stages are not clear. In the retina, fast adaptation in the receptive field center is thought to originate from synaptic mechanisms in bipolar cells (Demb, 2008). There is also a form of contrast gain control driven by remote stimulation far from the receptive field center of retinal ganglion cells. This adaptation is mediated by synaptic inhibition from long-range amacrine cells at the bipolar cell terminal (Demb, 2008). Our findings suggest the existence of a similar wide-field amacrine cell in the fly medulla that pools local responses over several columns in order to estimate average contrast in a part of the visual scene.

3.2.4 Future experiments

While we have described basic tuning properties of contrast-dependent suppression in the fly, there are still many open questions regarding the implementation of this mechanism.

Through blocking experiments, we could show that the suppressive signal partially originates from feedback connections projecting from medulla neurons back onto themselves. However, we tested only the possibility of a "private pool", i.e. direct feedback suppression on a given cell type through a pool comprising the same cell type. This was mainly due to experimental constraints, so it is entirely possible that in fact there is a "public pool" that comprises several cell types and is shared among several cell types. For example, silencing Tm1 had almost no effects on Tm1 responses suggesting that the suppressive signal that modulates Tm1 originates from other cell types. On the other hand, blocking Tm2 led to an almost complete loss of contrast-dependent response modulation in Tm2. Future experiments

should therefore test if silencing Tm2 also affects the strength of contrast gain control in other cell types. In general, there could be one or more pool cells, each integrating signals from one or more subtypes of medulla cells. Only a systematic approach scanning all combinations of silencing and imaging different cell types across all medulla cells can shed light on the underlying composition of the respective normalization signals. Specifically, it will be interesting to see if ON- and OFF-pathway cells share common sources of suppression or if they are functionally completely separated in this regard. There is evidence that contrast gain adaptation is asymmetric between ON- and OFF-channels in the retina (Chander and Chichilnisky, 2001; Zaghloul et al., 2005). In addition, we found generally different baseline contrast sensitivities between ON- and OFF-pathway elements in the fly, which suggests the possibility that normalization might be split into an ON- and an OFF-system.

The suppressive gain control signal that we describe is sensitive to dynamic contrast, has temporal band-pass tuning but is not direction selective. We found that a pool cell mechanism averaging appropriate columnar neurons that are not direction selective and exhibit temporal band-pass properties accounts well for these observations. Therefore, measuring the spatial frequency tuning of the suppressive signal should also reflect the spatial filtering characteristics of the input cells to the pool. Since most medulla neurons possess spatially antagonistic receptive fields with varying strength of surround suppression, I would predict a mild band-pass spatial frequency tuning for the normalization signal. The direction independence of the suppressive signal could in principle also result from a well-balanced pool of T4 and T5 cells including all four subtypes. To exclude this possibility, it should be confirmed that divisive suppression in medulla cells is indeed independent of motion. This could be achieved by using uncorrelated fluctuations as a background stimulus, similar to our behavioral stimulus, or by testing if blocking T4 or T5 cells affects normalization strength in the medulla.

Furthermore, it will be interesting to learn more about the dynamics of contrast normalization in the fly. The contrast-step protocol used in our experiments sampled the strength of suppression only every 50 ms at most. While this allowed us to resolve transient ringing after step onset in the suppressive signal, this temporal precision is too low to exactly quantify the short onset time constant of the suppression. Onset latency of cross-orientation suppression in the cortex was found to lie between 20–80ms (Smith et al., 2006). Therefore, a similar experiment should be repeated with significantly smaller time steps to characterize onset dynamics better. Additionally, it would be interesting to scan also the offset dynamics of contrast-dependent suppression with a similar stimulus protocol.

One question that we did not resolve in our study is whether surround suppression indeed acts in a divisive way also for the OFF-pathway elements Tm1, Tm2 and Tm4. In divisive contrast normalization, contrast gain is regulated such that local contrast tuning curves are shifted horizontally on a logarithmic axis, e.g. as observed for Mi1 and Tm3 cells in our dataset. For the OFF-pathway elements, however, responses did not yet saturate even for the highest contrast in our stimulus protocol. Therefore, it is impossible

to say if the effect of suppression by surround contrast is a horizontal or a vertical shift of the tuning curves. In cases like these, alternative models such as, for example, subtractive suppression can be fit equally well (Cavanaugh et al., 2002a). Using a more efficient driving stimulus in the center of the receptive field, e.g. local intensity steps, might lead to response saturation for these cell types and resolve this issue.

Finally, our study opens the field for circuit mapping in order to pinpoint the neuronal mechanisms of contrast normalization. Our data suggests the existence of a pool cell that computes average contrast over an area of approximately 50-70° in visual space. This putative neuron should therefore span its dendritic tree over approximately ten columns, although coupling between neighbouring cells could reduce the necessary diameter of the dendritic arbor. If we assume that the suppressive signal is not further relayed via another interneuron we can narrow down speculations about the morphology of that putative wide-field cell. For the ON-pathway, the neuron should reside completely within the medulla with its dendrites likely in layer M10 in order to make synaptic contact with Mi1 and Tm3 axon terminals. Feedback inhibition could be provided by sending neurites back to more distal medulla layers where Mi1 and Tm3 possess stratifications, but could also be exercised directly within layer M10. For OFF-pathway cells, axon terminals are located in the lobula, so the pool cell could be a lobula neuron. However, all normalized OFF-cells also have stratifications in at least two layers of the medulla. Hence, wide-field neurons in the medulla are also possible candidates for OFF-pathway normalization.

A number of wide-field amacrine cells in the fly optic lobe match such anatomical profile (Fischbach and Dittrich, 1989). Distal medulla (Dm) neurons are found in layer 1–6 of the medulla and span several columns. In addition to that, proximal medulla (Pm) neurons possess large dendritic arbors in layers M9 and M10 of the medulla and seem to perform local computations only there. For many of these cell types, Gal4-driver lines have already been made available (Nern et al., 2015; Davis et al., 2018). Transsynaptic mapping using the anterograde circuit tracing tool *trans-Tango* (Talay et al., 2017) under control of the UAS promoter might be a viable test for connectivity between candidate amacrine cells and normalized medulla neurons. Another approach would be to trace postsynaptic partners of normalized medulla neurons using electron microscopy. Existing datasets might contain some of the relevant connections already. However, the reconstructed volumes could be too small to allow for reliable identification of such large-field amacrine cells (Takemura et al., 2017; Shinomiya et al., 2019). Acquisition of new electron microscopy datasets and reconstruction of the anatomical structures contained in such volumes are large-scale projects that require joint efforts of a big research team. Ongoing connectomics projects of such kind are constantly improving the quality and the completeness of anatomical reconstructions of the *Drosophila* brain (see Janelia FlyEM Project). Therefore, it might be possible already in the near future to look for wide-field amacrine cells postsynaptic to normalized medulla neurons in a complete connectome of the fly optic lobe.

After identification of possible candidate neurons that implement the mechanism, ultimate proof for the validity of this "pool-cell hypothesis"

will require silencing or activation experiments. The prediction of our model is that blocking the output of such pool cell would lead to linearity of postsynaptic medulla neurons so that they respond independently of background contrast. Activation experiments could have two outcomes: either the resulting suppressive signal is so strong that affected medulla cells become dysfunctional; or the suppressive signal would stay at an elevated maximum level, though not strong enough to completely suppress responses in postsynaptic neurons. In the latter case, I would expect smaller but also linear responses in the affected medulla neurons, since, as in blocking experiments, the context-dependency induced by surround suppression would be lost.

3.3 PROJECTOR-BASED ARENA

Visual information processing is investigated by precise analysis of the output of a neural system given a certain input stimulus. Often, the assessment of the output of a system is heavily constrained by the availability and the accuracy of the experimental tools and techniques. Therefore, it is of utmost importance to minimize technical constraints when it comes to design of input stimuli. It was a crucial part of my doctoral work to design a visual stimulation system for experiments with fruit flies that allows for display of arbitrary stimuli with high spatial as well as temporal precision. The goal was to provide researchers with maximal control over the input stimulus during calcium imaging or electrophysiology experiments. In the following, advantages and limitations of my technical approach to the design of a projector-based stimulation arena for fruit flies will be outlined.

3.3.1 Visual stimulation for fruit flies

When building a visual stimulation device for flies one has to consider the specific requirements that are imposed by the biology of fly vision. For compound eyes, spatial resolution lies in the range of the divergence angle of neighboring photoreceptors, which is around 4.6° for *Drosophila* (Götz, 1964). Since this is a relatively poor spatial resolution compared to human vision, there is no need for use of high-resolution displays when stimulating fruit flies visually. In contrast to that, flies can resolve much finer temporal detail than humans. As pointed out above (see Section 1.2.2), fly photoreceptors can resolve temporal changes up to frequencies between 100-200 Hz. Since this lies beyond the refresh rate of most commercial displays, it poses one of the major challenges for visual stimulation systems for flies.

In early experiments, patterned cylinders were mechanically rotated around the fly to generate whole-field motion stimuli (Reichardt, 1961; Götz, 1964). While this approach does not face the problem of limited refresh rates, since it does not involve electronic displays, it is very constrained with respect to stimulus design.

To overcome this problem, LED-based stimulation arenas have been developed and used extensively in the field of fly motion vision (Reiser and

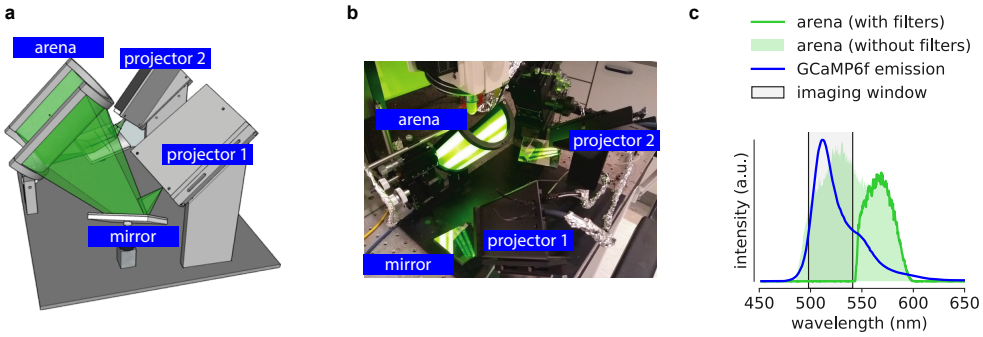


Figure 15: Projector-based arena design | **a**, 3D model of the arena design. Green shaded areas indicate light paths from the projectors to the rear projection screen. Graphics courtesy of S. Prech. **b**, Photo of the arena installed in an electrophysiology setup. **c**, Spectrum of the arena light. Light green shaded area shows spectrum without long-pass filters. Green solid line shows spectrum after long-pass filtering at 550 nm as used for calcium imaging, normalized to the peak of the light green area in the corresponding part of the spectrum. Blue solid line indicates GCaMP6f emission spectrum (data obtained from Harris Lab (2019)). Gray shaded area indicates the spectral window for detection of fluorescence signals (transmission window of a Semrock FF01-520/35 filter).

Dickinson, 2008; Joesch et al., 2008). These arenas consist of modular LED array panels which are arranged in a cylindrical fashion to maximize coverage of the fly visual field. Stimulation patterns are usually pre-rendered and uploaded onto the arena control board which is operated via PC. The size of the individual LEDs amounts up to a few degrees of viewing angle, which is enough considering fly spatial resolution limits. These stimulation systems can operate at refresh rates between 400-500 Hz, restricted only by the data transmission rate between the display controller and the LED panels (Reiser and Dickinson, 2008; Joesch et al., 2008). This is ideal for stimulation of fly photoreceptors and offers elegant ways to prevent bleed-through of visual stimuli into fluorescence signals during calcium imaging experiments by interleaving image acquisition and stimulus display. A drawback, however, is the necessity to pre-render stimuli. This can be impractical and forbids customization of stimuli during the experiment. In order to address this issue, I chose an approach which would allow for live-rendering of stimuli and therefore would give the experimenter full control over the visual stimulus during an experiment.

3.3.2 Design of the projector-based arena

To design a new stimulation arena for flies, I took advantage of micro-projector technology. The arena consists of a cylindrical aluminium frame on which a rear-projection foil is mounted (Figure 15a,b). The foil material was chosen such as to guarantee high transmittance and isotropy of scattered light (type *EVEN* by *Gerriets GmbH*). Scattering isotropy is of particular importance here because the incidence angle of the light beam from the projector depends heavily on the position on the screen. Two micro-projectors were placed on aluminium stages such that their projected images appeared exactly on the two halves of the cylindrical rear-projection screen. This light

path was established via two additional mirrors below the projectors to allow for a more compact design.

The micro-projector that is installed in the system is the *DLP LightCrafter 3000* by *Texas Instruments*. This projector relies on a digital micromirror device that consists of an array of 608 x 684 micromirrors, each one corresponding to one pixel in the projected image. The light intensity in each pixel is controlled through rapid tilting of the corresponding mirror which effectively sets the illumination time of each pixel individually. In RGB mode, projectors of this type present the three color channels of an image temporally interleaved. We modified the projectors to use only the green LED for all color channels which permits reassignment of each color channel as an additional temporal frame effectively increasing the frame rate from 60 Hz to 180 Hz. Separation of green stimulus light from fluorescence signals in calcium imaging applications is implemented by placing long-pass filters with a cutoff at 550 nm in front of the projectors. The resulting spectral distribution of the arena light is shown in Figure 15c.

Without modifications, projecting an image like this onto a curved screen would induce severe image distortions. Therefore, projected images are predistorted before display. This is achieved by taking advantage of GPU-based rendering: Replicating the geometry of the arena in a virtual model allows to generate a view of the virtual arena screen from the point in space where the projector is placed. The image taken under this perspective accounts for all distortions due to the geometry of the set-up. Re-projecting this image back onto the real arena screen results in regular display of the stimulus which is set as a texture on the virtual arena screen.

Stimuli are live-rendered on the GPU of the control computer. This approach facilitates parametrization of stimuli, offers maximum flexibility in stimulus design and, in principle, also enables closed-loop and interactive user-controlled stimuli.

3.3.3 Light diffusion properties of the arena

Any image projection system is based on scattering of photons on the screen material. In general, light will be scattered in all directions of space, more or less isotropically depending on the material. Although this is intentional because it also includes the direction of the observer, which in this case is a fly, there will also be light scattered into other regions on the screen, eventually causing optical artifacts like blur or reflections. This section provides a short characterization of light diffusion artifacts in the arena.

In order to measure light intensity in a point-like area on the screen, I attached a light guide onto the screen surface and connected it on the other side to a photodiode (Figure 16a). The resulting photocurrent is proportional to and a direct read-out of the screen light intensity at the position of the light guide. For confirmation, I reconstructed the receptive field of the light guide using white noise analysis. When using a white noise stimulus that tiled the screen into 128 pixels along the horizontal axis, I obtained a small circular receptive field exactly at the position of the light

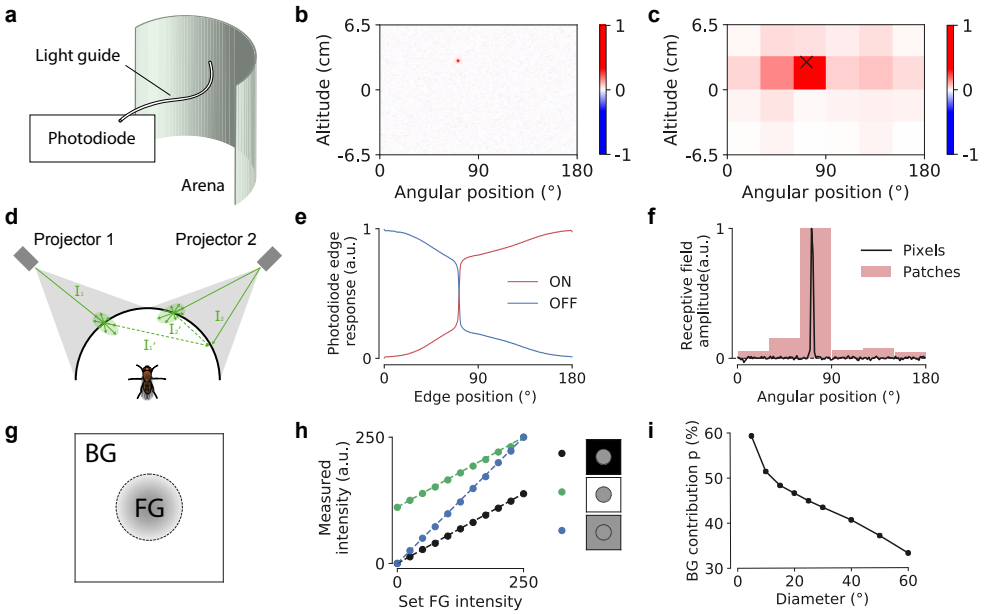


Figure 16: Photodiode measurements of local luminance | **a**, Schematic of the arena with light guide attached to the screen. Using a photodiode, the intensity of the light transmitted through the light guide is measured. **b**, Receptive field of the light guide using a white noise stimulus with 128 pixels along the horizontal axis. **c**, Receptive field of the light guide using a white noise stimulus partitioning the screen into 6 patches along the horizontal axis. The cross marks the receptive field position from the 128 pixels white noise. **d**, Schematic of light scattering within the arena. Light intensity at a given point in the arena is the sum of direct illumination and stray light from other areas on the screen. **e**, Edge response of the photodiode to ON (red) and OFF (blue) edges. **f**, Horizontal cross-sections through the receptive field from the pixels (black line) and the patches white noise (red bars). **g**, Schematic of the stimulus. A local foreground window (FG) is set to a different light intensity than the background (BG). The background signal leaks into the foreground due to light diffusion. **h**, Actual light intensity in the foreground versus set foreground intensity, for black background (black dots), white background (green dots), and background intensity equal to foreground intensity (blue dots). Data shown for a 25° diameter of the foreground window. Dashed lines show model fit. **i**, Background contribution p from the model fit in dependence of the diameter of the foreground window.

guide (Figure 16b). However, when partitioning the screen into 6 large patches along the horizontal, I observed significant correlations between photocurrent and light intensity not only locally but also from remote areas on the screen (Figure 16c,f). The reason for that could be stray light which is not detected by the analysis with the pixel noise stimulus because significant deviations from average intensity over large areas are extremely unlikely for such fine-grained noise stimuli. When considering the geometry of the arena, it becomes clear how light diffusion can modify the local light intensity at a spot: Due to the cylindrical shape of the screen, the light intensity at a given point is, in fact, the sum of direct illumination from the projectors and stray light contributions from neighboring points on both halves of the cylinder (Figure 16d). To demonstrate this more clearly, I measured the edge response of the photodiode. When moving light or dark edges horizontally across the

screen, responses are smoothed out considerably over large distances due to light diffusion (Figure 16e).

To quantify the amount of stray light leakage into local signals, I varied light intensity in the foreground, defined by a circular window around the light guide position, independently from the background, defined as the rest of the arena (Figure 16g). As expected, local luminance depended linearly on the set foreground light intensity, but there was a constant offset to the curve with background at maximum intensity compared to when the background was black (Figure 16h). When foreground and background intensity were stepped up together, measured light intensities lied along a diagonal connecting these two conditions. I fit a model to these curves assuming that stray light contributes linearly to local luminance. Specifically, I fit the equation

$$y = x_{fg} \cdot (1 - p) + x_{bg} \cdot p$$

where y is the measured local luminance, x_{fg} is the set foreground intensity, x_{bg} the set background intensity, and p the percentage to which the background contributes to local luminance through light diffusion. This linear model accounts well for the observed double-dependency of local luminance on set foreground and background intensity (Figure 16h). I repeated this experiment for different diameters of the foreground window and fit the model parameter p independently for each condition (Figure 16i). While for a diameter of 5° stray light accounts for more than 50% of the signal amplitude, this contribution drops rapidly with increasing diameter. Nonetheless, light diffusion contributes substantially to local light intensity even for large diameters.

In practice, this means that light diffusion in the arena should be taken into account when designing new stimuli and for the interpretation of experimental results. As a first approximation, it is safe to assume that for most stimuli light scattering will lead to a constant offset of the set light intensity in each pixel by some value according to the average luminance over the whole screen. In general, this will result in a net decrease of image contrast. Fine spatial structures of the stimulus might be blurred out because light diffusion is more severe over small distances. Since light scattering is a physical process that is hard to control, these are the limitations that should be considered for virtually all display systems to some degree.

One important point to take into account are adaptation processes in the early visual system that could induce non-linear distortions to neuronal responses. For example, an increase in light intensity due to stray light originating from an approaching bright edge might already be registered as a strong change in stimulus contrast by dark-adapted photoreceptors. A subsequent neuron might enhance these signals and respond strongly to the scattered light although the actual edge has not yet even passed the receptive field of the cell. Setting a background intensity which is not zero as a baseline for dark regions of the stimulus might help to alleviate effects like these and to linearize neuronal responses. On the other hand, effective contrast could also be intentionally boosted for local stimuli by setting the background intensity to zero. The experimenter should always be aware of these processes and take them into account when tailoring visual stimuli for a given scientific question.

3.4 CONCLUDING REMARKS

Taken together, the work that I presented in this cumulative dissertation advanced our knowledge about motion processing in the fly in several important directions. First, we systematically mapped the functional characteristics of the majority of synaptic input neurons to T4 and T5 cells and hence provided a rich resource for future computational approaches in the field. Second, we discovered that behavioral state modulation acts already in the early visual system in the fly. Third, we addressed related questions about how motion cues can be further processed in order to enable visually guided tracking of conspecific flies during male courtship behavior. Finally, we revealed a level of computational complexity of the early visual system in *Drosophila* that had not been attributed before to this seemingly straightforward circuitry. The discovery of feedback mechanisms that regulate the contrast sensitivity of medulla neurons opens the field to a wide range of related research questions. Which neurons represent the pool that computes local contrast? Which neurons feed into this pool? How is the suppressive mechanism implemented at the synaptic and molecular level? Are there intrinsic asymmetries between ON- and OFF-pathway regarding contrast normalization in the fly? It is indeed the very nature of science that by finding the solution to one problem we discover a whole new set of questions awaiting an answer in the future.

BIBLIOGRAPHY

Adelson, E. H. and J. R. Bergen
1985. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America. A, Optics and Image Science*, 2(2):284–299.

Ammer, G., A. Leonhardt, A. Bahl, B. J. Dickson, and A. Borst
2015. Functional specialization of neural input elements to the *Drosophila* ON motion detector. *Current Biology*, 25(17):2247–2253.

Arenz, A., M. S. Drews, F. G. Richter, G. Ammer, and A. Borst
2017. The temporal tuning of the *Drosophila* motion detectors is determined by the dynamics of their input elements. *Current Biology*, 27(7):929–944.

Attneave, F.
1954. Some informational aspects of visual perception. *Psychological Review*, 61(3):183–193.

Baccus, S. A. and M. Meister
2002. Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36(5):909–919.

Baden, T., P. Berens, M. Bethge, and T. Euler
2013. Spikes in mammalian bipolar cells support temporal layering of the inner retina. *Current Biology*, 23(1):48–52.

Baines, R. A., J. P. Uhler, A. Thompson, S. T. Sweeney, and M. Bate
2001. Altered electrical properties in *Drosophila* neurons developing without synaptic transmission. *Journal of Neuroscience*, 21(5):1523–1531.

Bair, W., J. R. Cavanaugh, and J. Anthony Movshon
2003. Time course and time-distance relationships for surround suppression in macaque V1 neurons. *Journal of Neuroscience*, 23(20):7690–7701.

Barlow, H. B.
1953. Summation and inhibition in the frog's retina. *Journal of Physiology*, 119(1):69–88.

Barlow, H. B.
1961. Possible principles underlying the transformation of sensory messages. *Sensory communication*, 1:217–234.

Barlow, H. B. and W. R. Levick
1965. The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, 178(3):477–504.

Bath, D. E., J. R. Stowers, D. Hörmann, A. Poehlmann, B. J. Dickson, and A. D. Straw
2014. FlyMAD: rapid thermogenetic control of neuronal activity in freely walking *Drosophila*. *Nature Methods*, 11(7):756–762.

Behnia, R., D. A. Clark, A. G. Carter, T. R. Clandinin, and C. Desplan
2014. Processing properties of ON and OFF pathways for *Drosophila* motion detection. *Nature*, 512(7515):427–430.

Bell, A. J. and T. J. Sejnowski
1997. The "independent components" of natural scenes are edge filters. *Vision Research*, 37(23):3327–3338.

Bellen, H. J., C. Tong, and H. Tsuda
2010. 100 years of *Drosophila* research and its impact on vertebrate neuroscience: a history lesson for the future. *Nature Reviews Neuroscience*, 11(7):514–522.

Berni, J., A. M. Muldal, and S. R. Pulver
2010. Using neurogenetics and the warmth-gated ion channel TRPA1 to study the neural basis of behavior in *Drosophila*. *Journal of Undergraduate Neuroscience Education*, 9(1):A5–A14.

Bonin, V., V. Mante, and M. Carandini
2005. The suppressive field of neurons in lateral geniculate nucleus. *Journal of Neuroscience*, 25(47):10844–10856.

Bonin, V., V. Mante, and M. Carandini
2006. The statistical computation underlying contrast gain control. *Journal of Neuroscience*, 26(23):6346–6353.

Borst, A.
2007. Correlation versus gradient type motion detectors: the pros and cons. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1479):369–374.

Borst, A.
2014. Fly visual course control: behaviour, algorithms and circuits. *Nature Reviews Neuroscience*, 15(9):590–599.

Borst, A.
2018. A biophysical mechanism for preferred direction enhancement in fly motion vision. *PLoS computational biology*, 14(6):e1006240.

Borst, A. and H. D. I. Abarbanel
2007. Relating a calcium indicator signal to the unperturbed calcium concentration time-course. *Theoretical Biology and Medical Modelling*, 4:7.

Borst, A. and M. Egelhaaf
1989. Principles of visual motion detection. *Trends in neurosciences*, 12(8):297–306.

Borst, A., M. Egelhaaf, and J. Haag
1995. Mechanisms of dendritic integration underlying gain control in fly motion-sensitive interneurons. *Journal of computational neuroscience*, 2(1):5–18.

Borst, A., V. L. Flanagin, and H. Sompolinsky
2005. Adaptation without parameter change: dynamic gain control in motion detection. *Proceedings of the National Academy of Sciences* , 102(17):6172–6176.

Borst, A., J. Haag, and D. F. Reiff
2010. Fly motion vision. *Annual Review of Neuroscience*, 33:49–70.

Borst, A. and M. Helmstaedter
2015. Common circuit design in fly and mammalian motion vision. *Nature Neuroscience*, 18(8):1067–1076.

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

Brady, N. and D. J. Field
2000. Local contrast in natural images: normalisation and coding efficiency. *Perception*, 29(9):1041–1055.

Brand, A. H. and N. Perrimon
1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, 118(2):401–415.

Briggman, K. L., M. Helmstaedter, and W. Denk
2011. Wiring specificity in the direction-selectivity circuit of the retina. *Nature*, 471(7337):183–188.

Brinkworth, R. S. A. and D. C. O'Carroll
2009. Robust models for optic flow coding in natural scenes inspired by insect biology. *PLoS computational biology*, 5(11):e1000555.

Busch, S., M. Selcho, K. Ito, and H. Tanimoto
2009. A map of octopaminergic neurons in the *Drosophila* brain. *Journal of Comparative Neurology*, 513(6):643–667.

Busse, L., A. R. Wade, and M. Carandini
2009. Representation of concurrent stimuli by population activity in visual cortex. *Neuron*, 64(6):931–942.

Carandini, M. and D. J. Heeger
2011. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62.

Carandini, M., D. J. Heeger, and J. Anthony Movshon
1997. Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience*, 17(21):8621–8644.

Cavanaugh, J. R., W. Bair, and J. A. Movshon
2002a. Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5):2530–2546.

Cavanaugh, J. R., W. Bair, and J. A. Movshon
2002b. Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5):2547–2556.

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher
1994. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805.

Chander, D. and E. J. Chichilnisky
2001. Adaptation to temporal contrast in primate and salamander retina. *Journal of Neuroscience*, 21(24):9904–9916.

Chen, T.-W., T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, and D. S. Kim
2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*, 499(7458):295–300.

Chiappe, M. E., J. D. Seelig, M. B. Reiser, and V. Jayaraman
2010. Walking modulates speed sensitivity in *Drosophila* motion vision. *Current Biology*, 20(16):1470–1475.

Clark, D. A., L. Bursztyn, M. A. Horowitz, M. J. Schnitzer, and T. R. Clandinin
2011. Defining the computational structure of the motion detector in *Drosophila*. *Neuron*, 70(6):1165–1177.

Cleland, T. A.
1996. Inhibitory glutamate receptor channels. *Molecular neurobiology*, 13(2):97–136.

Cosens, D. and H. C. Spatz
1978. Flicker fusion studies in the lamina and receptor region of the *Drosophila* eye. *Journal of Insect Physiology*, 24(8):587–594.

Dan, Y., J. J. Atick, and R. C. Reid
1996. Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *Journal of Neuroscience*, 16(10):3351–3362.

Davis, F. P., A. Nern, S. Picard, M. B. Reiser, G. M. Rubin, S. R. Eddy, and G. L. Henry
2018. A genetic, genomic, and computational resource for exploring neural circuit function. *bioRxiv*, 385476.

Dayan, P., G. E. Hinton, R. M. Neal, and R. S. Zemel
1995. The helmholtz machine. *Neural computation*, 7(5):889–904.

DeAngelis, G. C., I. Ohzawa, and R. D. Freeman
1995. Receptive-field dynamics in the central visual pathways. *Trends in neurosciences*, 18(10):451–458.

DeAngelis, G. C., J. G. Robson, I. Ohzawa, and R. D. Freeman
1992. Organization of suppression in receptive fields of neurons in cat visual cortex. *Journal of Neurophysiology*, 68(1):144–163.

Demb, J. B.
2008. Functional circuitry of visual adaptation in the retina. *Journal of Physiology*, 586(18):4377–4384.

Denk, W., J. H. Strickler, and W. W. Webb
1990. Two-photon laser scanning fluorescence microscopy. *Science*, 248(4951):73–76.

Ding, H., R. G. Smith, A. Poleg-Polsky, J. S. Diamond, and K. L. Briggman
2016. Species-specific wiring for direction selectivity in the mammalian retina. *Nature*, 535(7610):105–110.

Dror, R. O., D. C. O'Carroll, and S. B. Laughlin
2001. Accuracy of velocity estimation by reichardt correlators. *Journal of the Optical Society of America. A, Optics and Image Science*, 18(2):241–252.

Dubs, A.
1982. The spatial integration of signals in the retina and lamina of the fly compound eye under different conditions of luminance. *Journal of Comparative Physiology*, 146(3):321–343.

Dweck, H. K. M., S. A. M. Ebrahim, M. Thoma, A. A. M. Mohamed, I. W. Keesey, F. Trona, S. Lavista-Llanos, A. Svatoš, S. Sachse, M. Knaden, and B. S. Hansson
2015. Pheromones mediating copulation and attraction in *Drosophila*. *Proceedings of the National Academy of Sciences* , 112(21):E2829–35.

Dyakova, O. and K. Nordström
2017. Image statistics and their processing in insect vision. *Current opinion in insect science*, 24:7–14.

Egelhaaf, M.
1985. On the neuronal basis of figure-ground discrimination by relative motion in the visual system of the fly. *Biological Cybernetics*, 52(2):123–140.

Egelhaaf, M., A. Borst, and W. Reichardt
1989. Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. *Journal of the Optical Society of America. A, Optics and Image Science*, 6(7):1070–1087.

Fennema, C. L. and W. B. Thompson
1979. Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9(4):301–315.

Field, D. J.
1987. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America. A, Optics and Image Science*, 4(12):2379–2394.

Fischbach, K.-F. and A. P. M. Dittrich
1989. The optic lobe of *Drosophila melanogaster*. I. A Golgi analysis of wild-type structure. *Cell and tissue research*, 258(3):441–475.

Fisher, Y. E., J. C. S. Leong, K. Sporar, M. D. Ketkar, D. M. Gohl, T. R. Clandinin, and M. Silies
2015a. A class of visual neurons with wide-field properties is required for local motion detection. *Current Biology*, 25(24):3178–3189.

Fisher, Y. E., M. Silies, and T. R. Clandinin
2015b. Orientation selectivity sharpens motion detection in *Drosophila*. *Neuron*, 88(2):390–402.

Fitzgerald, J. E. and D. A. Clark
2015. Nonlinear circuits for naturalistic visual motion estimation. *Elife*, 4:e09123.

Franz, M. O. and H. G. Krapp
2000. Wide-field, motion-sensitive neurons and matched filters for optic flow fields. *Biological Cybernetics*, 83(3):185–197.

Frazor, R. A. and W. S. Geisler
2006. Local luminance and contrast in natural images. *Vision Research*, 46(10):1585–1598.

Freeman, T. C. B., S. Durand, D. C. Kiper, and M. Carandini
2002. Suppression without inhibition in visual cortex. *Neuron*, 35(4):759–771.

Freifeld, L., D. A. Clark, M. J. Schnitzer, M. A. Horowitz, and T. R. Clandinin
2013. GABAergic lateral interactions tune the early stages of visual processing in *Drosophila*. *Neuron*, 78(6):1075–1089.

Fujiwara, T., T. L. Cruz, J. P. Bohnslav, and M. E. Chiappe
2017. A faithful internal representation of walking movements in the *Drosophila* visual system. *Nature Neuroscience*, 20(1):72–81.

Geisler, W. S.
2008. Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, 59:167–192.

Geisler, W. S., J. S. Perry, B. J. Super, and D. P. Gallogly
2001. Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41(6):711–724.

Gilbert, C. and N. J. Strausfeld
1991. The functional organization of male-specific visual neurons in flies. *Journal of Comparative Physiology A*, 169(4):395–411.

Göppert-Mayer, M.
1931. Über Elementarakte mit zwei Quantensprüngen. *Annalen der Physik*, 401(3):273–294.

Götz, K. G.
1964. Optomotorische Untersuchung des visuellen Systems einiger Augenmutanten der Fruchtfliege *Drosophila*. *Kybernetik*, 2(2):77–92.

Götz, K. G.
1968. Flight control in *Drosophila* by visual perception of motion. *Kybernetik*, 4(6):199–208.

Götz, K. G., B. Hengstenberg, and R. Biesinger
1979. Optomotor control of wing beat and body posture in *Drosophila*. *Biological Cybernetics*, 35(2):101–112.

Govorunova, E. G., O. A. Sineshchekov, R. Janz, X. Liu, and J. L. Spudich
2015. Natural light-gated anion channels: A family of microbial rhodopsins for advanced optogenetics. *Science*, 349(6248):647–650.

Gruntman, E., S. Romani, and M. B. Reiser
2018. Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in *Drosophila*. *Nature Neuroscience*, 21(2):250–257.

Haag, J., A. Arenz, E. Serbe, F. Gabbiani, and A. Borst
2016. Complementary mechanisms create direction selectivity in the fly. *Elife*, 5:e17421.

Haag, J. and A. Borst
1998. Active membrane properties and signal encoding in graded potential neurons. *Journal of Neuroscience*, 18(19):7972–7986.

Haag, J. and A. Borst
2001. Recurrent network interactions underlying flow-field selectivity of visual interneurons. *Journal of Neuroscience*, 21(15):5685–5692.

Haag, J., W. Denk, and A. Borst
2004. Fly motion vision is based on Reichardt detectors regardless of the signal-to-noise ratio. *Proceedings of the National Academy of Sciences* , 101(46):16333–16338.

Haag, J., M. Egelhaaf, and A. Borst
1992. Dendritic integration of motion information in visual interneurons of the blowfly. *Neurosci. Lett.*, 140(2):173–176.

Haag, J., A. Mishra, and A. Borst
2017. A common directional tuning mechanism of *Drosophila* motion-sensing neurons in the ON and in the OFF pathway. *Elife*, 6:e29044.

Haikala, V., M. Joesch, A. Borst, and A. S. Mauss
2013. Optogenetic control of fly optomotor responses. *Journal of Neuroscience*, 33(34):13927–13934.

Hall, J. C.
1994. The mating of a fly. *Science*, 264(5166):1702–1714.

Hamada, F. N., M. Rosenzweig, K. Kang, S. R. Pulver, A. Ghezzi, T. J. Jegla, and P. A. Garrity
2008. An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature*, 454(7201):217–220.

Hancock, P. J. B., R. J. Baddeley, and L. S. Smith
1992. The principal components of natural images. *Network: Computation in Neural Systems*, 3(1):61–70.

Hardie, R. C.
1989. A histamine-activated chloride channel involved in neurotransmission at a photoreceptor synapse. *Nature*, 339(6227):704–706.

Hardie, R. C. and M. Juusola
2015. Phototransduction in *Drosophila*. *Current opinion in neurobiology*, 34:37–45.

Harris, R. A., D. C. O'Carroll, and S. B. Laughlin
2000. Contrast gain reduction in fly motion adaptation. *Neuron*, 28(2):595–606.

Harris Lab
2019. Two-photon fluorescent probes. (2019, May 16). Retrieved from *https://www.janelia.org/lab/harris-lab/research/photophysics/two-photon-fluorescent-probes*.

Hartline, H. K.
1940. The receptive fields of optic nerve fibers. *American Journal of Physiology-Legacy Content*, 130(4):690–699.

Harz, H. and P. Hegemann
1991. Rhodopsin-regulated calcium currents in *Chlamydomonas*. *Nature*, 351(6326):489–491.

Hassenstein, B.
1951. Ommatidienraster und afferente Bewegungsintegration. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 33(4):301–326.

Hassenstein, B. and W. Reichardt
1956. Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift für Naturforschung B*, 11(9-10):513–524.

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.

Hausen, K.
1982. Motion sensitive interneurons in the optomotor system of the fly. *Biological Cybernetics*, 45(2):143–156.

Hausen, K. and C. Wehrhahn
1983. Microsurgical lesion of horizontal cells changes optomotor yaw responses in the blowfly *Calliphora erythrocephala*. *Proceedings of the Royal society of London. Series B. Biological sciences*, 219(1215):211–216.

Hayashi, S., K. Ito, Y. Sado, M. Taniguchi, A. Akimoto, H. Takeuchi, T. Aigaki, F. Matsuzaki, H. Nakagoshi, T. Tanimura, R. Ueda, T. Uemura, M. Yoshihara, and S. Goto
2002. GETDB, a database compiling expression patterns and molecular locations of a collection of Gal4 enhancer traps. *Genesis*, 34(1-2):58–61.

Hecht, S. and S. Shlaer
1936. Intermittent stimulation by light: V. the relation between intensity and critical frequency for different parts of the spectrum. *The Journal of general physiology*, 19(6):965–977.

Heeger, D. J.
1992. Normalization of cell responses in cat striate cortex. *Visual neuroscience*, 9(2):181–197.

Honegger, H.-W.
1980. Receptive fields of sustained medulla neurons in crickets. *Journal of Comparative Physiology*, 136(3):191–201.

Hopp, E., A. Borst, and J. Haag
2014. Subcellular mapping of dendritic activity in optic flow processing neurons. *Journal of Comparative Physiology A*, 200(5):359–370.

Hubel, D. H. and T. N. Wiesel
1959. Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology*, 148:574–591.

Hyvärinen, A., J. Hurri, and P. O. Hoyer
2009. *Natural Image Statistics*, volume 39 of *A Probabilistic Approach to Early Computational Vision.* London: Springer Science & Business Media.

Jenett, A., G. M. Rubin, T.-T. B. Ngo, D. Shepherd, C. Murphy, H. Dionne, B. D. Pfeiffer, A. Cavallaro, D. Hall, J. Jeter, N. Iyer, D. Fetter, J. H. Hausenfluck, H. Peng, E. T. Trautman, R. R. Svirskas, E. W. Myers, Z. R. Iwinski, Y. Aso, G. M. DePasquale, A. Enos, P. Hulamm, S. C. B. Lam, H.-H. Li, T. R. Laverty, F. Long, L. Qu, S. D. Murphy, K. Rokicki, T. Safford, K. Shaw, J. H. Simpson, A. Sowell, S. Tae, Y. Yu, and C. T. Zugates
2012. A GAL4-driver line resource for *Drosophila* neurobiology. *Cell reports*, 2(4):991–1001.

Joesch, M., J. Plett, A. Borst, and D. F. Reiff
2008. Response properties of motion-sensitive visual interneurons in the lobula plate of *Drosophila melanogaster*. *Current Biology*, 18(5):368–374.

Joesch, M., B. Schnell, S. V. Raghu, D. F. Reiff, and A. Borst
2010. ON and OFF pathways in *Drosophila* motion vision. *Nature*, 468(7321):300–304.

Johns, D. C., R. Marx, R. E. Mains, B. O'Rourke, and E. Marbán
1999. Inducible genetic suppression of neuronal excitability. *Journal of Neuroscience*, 19(5):1691–1697.

Jung, S. N., A. Borst, and J. Haag
2011. Flight activity alters velocity tuning of fly motion-sensitive neurons. *Journal of Neuroscience*, 31(25):9231–9237.

Kim, A. J., L. M. Fenk, C. Lyu, and G. Maimon
2017. Quantitative predictions orchestrate visual signaling in *Drosophila*. *Cell*, 168(1-2):280–294.e12.

Kim, A. J., J. K. Fitzgerald, and G. Maimon
2015. Cellular evidence for efference copy in *Drosophila* visuomotor processing. *Nature Neuroscience*, 18(9):1247–1255.

Kirschfeld, K.
1967. Die Projektion der optischen Umwelt auf das Raster der Rhabdomere im Komplexauge von *Musca*. *Experimental Brain Research*, 3(3):248–270.

Kitamoto, T.
2001. Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive shibire allele in defined neurons. *Journal of Neurobiology*, 47(2):81–92.

Klapoetke, N. C., Y. Murata, S. S. Kim, S. R. Pulver, A. Birdsey-Benson, Y. K. Cho, T. K. Morimoto, A. S. Chuong, E. J. Carpenter, Z. Tian, J. Wang, Y. Xie, Z. Yan, Y. Zhang, B. Y. Chow, B. Surek, M. Melkonian, V. Jayaraman, M. Constantine-Paton, G. K.-S. Wong, and E. S. Boyden
2014. Independent optical excitation of distinct neural populations. *Nature Methods*, 11(3):338–346.

Kolodziejczyk, A., X. Sun, I. A. Meinertzhagen, and D. R. Nässel
2008. Glutamate, GABA and acetylcholine signaling components in the lamina of the *Drosophila* visual system. *PLoS One*, 3(5):e2110.

Kosaka, T. and K. Ikeda
1983. Possible temperature-dependent blockage of synaptic vesicle recycling induced by a single gene mutation in *Drosophila*. *Journal of Neurobiology*, 14(3):207–225.

Krapp, H. G., B. Hengstenberg, and R. Hengstenberg
1998. Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. *Journal of Neurophysiology*, 79(4):1902–1917.

Krapp, H. G. and R. Hengstenberg
1996. Estimation of self-motion by optic flow processing in single visual interneurons. *Nature*, 384(6608):463–466.

Kuffler, S. W.
1953. Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16(1):37–68.

Land, M. F. and T. S. Collett
1974. Chasing behaviour of houseflies (*Fannia canicularis*). *Journal of Comparative Physiology*, 89(4):331–357.

Lee, T. and L. Luo
1999. Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron*, 22(3):451–461.

Leong, J. C. S., J. J. Esch, B. Poole, S. Ganguli, and T. R. Clandinin
2016. Direction selectivity in *Drosophila* emerges from preferred-direction enhancement and null-direction suppression. *Journal of Neuroscience*, 36(31):8078–8092.

Leonhardt, A., G. Ammer, M. Meier, E. Serbe, A. Bahl, and A. Borst
2016. Asymmetry of *Drosophila* ON and OFF motion detectors enhances real-world velocity estimation. *Nature Neuroscience*, 19(5):706–715.

Li, B., J. K. Thompson, T. Duong, M. R. Peterson, and R. D. Freeman
2006. Origins of cross-orientation suppression in the visual cortex. *Journal of Neurophysiology*, 96(4):1755–1764.

Lima, S. Q. and G. Miesenböck
2005. Remote control of behavior through genetically targeted photostimulation of neurons. *Cell*, 121(1):141–152.

Limb, J. O. and J. A. Murphy
1975. Estimating the velocity of moving images in television signals. *Computer Graphics and Image Processing*, 4(4):311–327.

Lin, J. Y., P. M. Knutsen, A. Muller, D. Kleinfeld, and R. Y. Tsien
2013. ReaChR: a red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation. *Nature Neuroscience*, 16(10):1499–1508.

Lin, M. Z. and M. J. Schnitzer
2016. Genetically encoded indicators of neuronal activity. *Nature Neuroscience*, 19(9):1142–1153.

Lindeberg, T.
2013. A computational theory of visual receptive fields. *Biological Cybernetics*, 107(6):589–635.

Liu, W. W. and R. I. Wilson
2013. Glutamate is an inhibitory neurotransmitter in the *Drosophila* olfactory system. *Proceedings of the National Academy of Sciences* , 110(25):10294–10299.

Luan, H., N. C. Peabody, C. R. Vinson, and B. H. White
2006. Refined spatial manipulation of neuronal function by combinatorial restriction of transgene expression. *Neuron*, 52(3):425–436.

Luo, S. X., R. Axel, and L. F. Abbott
2010. Generating sparse and selective third-order responses in the olfactory system of the fly. *Proceedings of the National Academy of Sciences* , 107(23):10713–10718.

Maddess, T., S. B. Laughlin, G. A. Horridge, and W. R. Levick
1985. Adaptation of the motion-sensitive neuron H1 is generated locally and governed by contrast frequency. *Proceedings of the Royal Society of London. Series B. Biological Sciences*.

Maimon, G., A. D. Straw, and M. H. Dickinson
2008. A simple vision-based algorithm for decision making in flying *Drosophila*. *Current Biology*, 18(6):464–470.

Maimon, G., A. D. Straw, and M. H. Dickinson
2010. Active flight increases the gain of visual motion processing in *Drosophila*. *Nature Neuroscience*, 13(3):393–399.

Maisak, M. S., J. Haag, G. Ammer, E. Serbe, M. Meier, A. Leonhardt, T. Schilling, A. Bahl, G. M. Rubin, A. Nern, B. J. Dickson, D. F. Reiff, E. Hopp, and A. Borst
2013. A directional tuning map of *Drosophila* elementary motion detectors. *Nature*, 500(7461):212–216.

Mank, M. and O. Griesbeck
2008. Genetically encoded calcium indicators. *Chemical reviews*, 108(5):1550–1564.

Marvin, J. S., B. G. Borghuis, L. Tian, J. Cichon, M. T. Harnett, J. Akerboom, A. Gordus, S. L. Renninger, T.-W. Chen, C. I. Bargmann, M. B. Orger, E. R. Schreiter, J. B. Demb, W.-B. Gan, S. A. Hires, and L. L. Looger
2013. An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nature Methods*, 10(2):162–170.

Mauss, A. S., M. Meier, E. Serbe, and A. Borst
2014. Optogenetic and pharmacologic dissection of feedforward inhibition in *Drosophila* motion vision. *Journal of Neuroscience*, 34(6):2254–2263.

Mauss, A. S., K. Pankova, A. Arenz, A. Nern, G. M. Rubin, and A. Borst
2015. Neural circuit to integrate opposing motions in the visual field. *Cell*, 162(2):351–362.

Mauss, A. S., A. Vlasits, A. Borst, and M. Feller
2017. Visual circuits for direction selectivity. *Annual Review of Neuroscience*, 40:211–230.

Meier, M. and A. Borst
2019. Extreme compartmentalization in a *Drosophila* amacrine cell. *Current Biology*.

Meier, M., E. Serbe, M. S. Maisak, J. Haag, B. J. Dickson, and A. Borst
2014. Neural circuit components of the *Drosophila* OFF motion vision pathway. *Current Biology*, 24(4):385–392.

Miall, R. C.
1978. The flicker fusion frequencies of six laboratory insects, and the response of the compound eye to mains fluorescent 'ripple'. *Physiological Entomology*, 3(2):99–106.

Montell, C.
1999. Visual transduction in *Drosophila*. *Annual Review of Cell and Developmental Biology*, 15:231–268.

Morgan, T. H.
1910. Sex limited inheritance in *Drosophila*. *Science*, 32(812):120–122.

Muijres, F. T., M. J. Elzinga, J. M. Melis, and M. H. Dickinson
2014. Flies evade looming targets by executing rapid visually directed banked turns. *Science*, 344(6180):172–177.

Nagel, G., M. Brauner, J. F. Liewald, N. Adeishvili, E. Bamberg, and A. Gottschalk
2005. Light activation of Channelrhodopsin-2 in excitable cells of *Caenorhabditis elegans* triggers rapid behavioral responses. *Current Biology*, 15(24):2279–2284.

Nern, A., B. D. Pfeiffer, and G. M. Rubin
2015. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proceedings of the National Academy of Sciences* , 112(22):E2967–76.

Nojima, T., A. Rings, and S. F. Goodwin
2018. Drosophila courtship: Love is not blind. *Current Biology*, 28(15):R840–R842.

Nordström, K., P. D. Barnett, and D. C. O'Carroll
2006. Insect detection of small targets moving in visual clutter. *PLoS Biology*, 4(3):e54.

Olsen, S. R., V. Bhandawat, and R. I. Wilson
2010. Divisive normalization in olfactory population codes. *Neuron*, 66(2):287–299.

Pankova, K. and A. Borst
2016. RNA-seq transcriptome analysis of direction-selective T4/T5 neurons in *Drosophila*. *PLoS One*, 11(9):e0163986.

Pankova, K. and A. Borst
2017. Transgenic line for the identification of cholinergic release sites in *Drosophila* melanogaster. *Journal of Experimental Biology*, 220(Pt 8):1405–1410.

Pearson, K.
1901. LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572.

Pfeiffer, B. D., A. Jenett, A. S. Hammonds, T.-T. B. Ngo, S. Misra, C. Murphy, A. Scully, J. W. Carlson, K. H. Wan, T. R. Laverty, C. Mungall, R. Svirskas, J. T. Kadonaga, C. Q. Doe, M. B. Eisen, S. E. Celniker, and G. M. Rubin
2008. Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proceedings of the National Academy of Sciences* , 105(28):9715–9720.

Pitkow, X. and M. Meister
2012. Decorrelation and efficient coding by retinal ganglion cells. *Nature Neuroscience*, 15(4):628–635.

Poggio, T. and W. Reichardt
1976. Visual control of orientation behaviour in the fly. Part II. Towards the underlying neural interactions. *Quarterly Reviews of Biophysics*, 9(3):377–438.

Potters, M. and W. Bialek
1994. Statistical mechanics and visual signal processing. *Journal de Physique I*, 4(11):1755–1775.

Rabinowitz, N. C., B. D. B. Willmore, J. W. H. Schnupp, and A. J. King
2011. Contrast gain control in auditory cortex. *Neuron*, 70(6):1178–1191.

Ramachandran, V. S.
1988. Perception of shape from shading. *Nature*, 331(6152):163–166.

Reichardt, W.
1961. Autocorrelation, a principle for evaluation of sensory information by the central nervous system. In *Sensory Communication*, W. A. Rosenblith, ed., Pp. 303–317. Cambridge: MIT Press, John Wiley.

Reichardt, W., T. Poggio, and K. Hausen
1983. Figure-ground discrimination by relative movement in the visual system of the fly. *Biological Cybernetics*, 46(1):1–30.

Reiff, D. F., J. Plett, M. Mank, O. Griesbeck, and A. Borst
2010. Visualizing retinotopic half-wave rectified input to the motion detection circuitry of *Drosophila*. *Nature Neuroscience*, 13(8):973–978.

Reiser, M. B. and M. H. Dickinson
2008. A modular display system for insect behavioral neuroscience. *Journal of Neuroscience Methods*, 167(2):127–139.

Reutsky-Gefen, I., L. Golan, N. Farah, A. Schejter, L. Tsur, I. Brosh, and S. Shoham
2013. Holographic optogenetic stimulation of patterned neuronal activity for vision restoration. *Nature communications*, 4:1509.

Ribeiro, I. M. A., M. Drews, A. Bahl, C. Machacek, A. Borst, and B. J. Dickson
2018. Visual projection neurons mediating directed courtship in *Drosophila*. *Cell*, 174(3):607–621.e18.

Richter, F. G., S. Fendl, J. Haag, M. S. Drews, and A. Borst
2018. Glutamate signaling in the fly visual system. *iScience*, 7:85–95.

Rivera-Alba, M., S. N. Vitaladevuni, Y. Mishchenko, Z. Lu, S.-Y. Takemura, L. Scheffer, I. A. Meinertzhagen, D. B. Chklovskii, and G. G. de Polavieja
2011. Wiring economy and volume exclusion determine neuronal placement in the *Drosophila* brain. *Current Biology*, 21(23):2000–2005.

Robie, A. A., A. D. Straw, and M. H. Dickinson
2010. Object preference by walking fruit flies, *Drosophila* melanogaster, is mediated by vision and graviperception. *Journal of Experimental Biology*, 213(Pt 14):2494–2506.

Rodieck, R. W.
1965. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, 5(11):583–601.

Rosenzweig, M., K. M. Brennan, T. D. Tayler, P. O. Phelps, A. Patapoutian, and P. A. Garrity
2005. The *Drosophila* ortholog of vertebrate TRPA1 regulates thermotaxis. *Genes & Development*, 19(4):419–424.

Ruderman, D. L.
1994. The statistics of natural images. *Network: Computation in Neural Systems*, 5(4):517–548.

Ruderman, D. L.
1997. Origins of scaling in natural images. *Vision Research*, 37(23):3385–3398.

Sakmann, B. and E. Neher
1984. Patch clamp techniques for studying ionic channels in excitable membranes. *Annual Review of Physiology*, 46:455–472.

Salazar-Gatzimas, E., M. Agrochao, J. E. Fitzgerald, and D. A. Clark
2018. The neuronal basis of an illusory motion percept is explained by decorrelation of parallel motion pathways. *Current Biology*, 28(23):3748–3762.e8.

Salazar-Gatzimas, E., J. Chen, M. S. Creamer, O. Mano, H. B. Mandel, C. A. Matulis, J. Pottackal, and D. A. Clark
2016. Direct measurement of correlation responses in *Drosophila* elementary motion detectors reveals fast timescale tuning. *Neuron*, 92(1):227–239.

Schilling, T. and A. Borst
2015. Local motion detectors are required for the computation of expansion flow-fields. *Biology Open*, 4(9):1105–1108.

Schwartz, O. and E. P. Simoncelli
2001. Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4(8):819–825.

Serbe, E., M. Meier, A. Leonhardt, and A. Borst
2016. Comprehensive characterization of the major presynaptic elements to the *Drosophila* OFF motion detector. *Neuron*, 89(4):829–841.

Shaner, N. C., R. E. Campbell, P. A. Steinbach, B. N. Giepmans, A. E. Palmer, and R. Y. Tsien
2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnology*, 22(12):1567.

Shannon, C. E.
1949. Communication theory of secrecy systems. *Bell system technical journal*, 28(4):656–715.

Shapley, R. M. and J. D. Victor
1978. The effect of contrast on the transfer properties of cat retinal ganglion cells. *Journal of Physiology*, 285:275–298.

Shinomiya, K., G. Huang, Z. Lu, T. Parag, C. S. Xu, R. Aniceto, N. Ansari, N. Cheatham, S. Lauchie, E. Neace, O. Ogundeyi, C. Ordish, D. Peel, A. Shinomiya, C. Smith, S. Takemura, I. Talebi, P. K. Rivlin, A. Nern, L. K. Scheffer, S. M. Plaza, and I. A. Meinertzhagen
2019. Comparisons between the ON- and OFF-edge motion pathways in the *Drosophila* brain. *eLife*, 8:e40025.

Shinomiya, K., T. Karuppudurai, T.-Y. Lin, Z. Lu, C.-H. Lee, and I. A. Meinertzhagen
2014. Candidate neural substrates for off-edge motion detection in *Drosophila*. *Current Biology*, 24(10):1062–1070.

Shoemaker, P. A., D. C. O'carroll, and A. D. Straw
2005. Velocity constancy and models for wide-field visual motion detection in insects. *Biological cybernetics*, 93(4):275–287.

Silies, M., D. M. Gohl, Y. E. Fisher, L. Freifeld, D. A. Clark, and T. R. Clandinin
2013. Modular use of peripheral input channels tunes motion-detecting circuitry. *Neuron*, 79(1):111–127.

Simoncelli, E. P. and B. A. Olshausen
2001. Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24:1193–1216.

Simpson, J. H.
2009. Mapping and manipulating neural circuits in the fly brain. *Advances in Genetics*, 65:79–143.

Sinakevitch, I. and N. J. Strausfeld
2006. Comparison of octopamine-like immunoreactivity in the brains of the fruit fly and blow fly. *Journal of Comparative Neurology*, 494(3):460–475.

Smith, M. A., W. Bair, and J. A. Movshon
2006. Dynamics of suppression in macaque primary visual cortex. *Journal of Neuroscience*, 26(18):4826–4834.

Solomon, S. G., J. W. Peirce, N. T. Dhruv, and P. Lennie
2004. Profound contrast adaptation early in the visual pathway. *Neuron*, 42(1):155–162.

Srinivasan, M. V., S. B. Laughlin, and A. Dubs
1982. Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 216(1205):427–459.

Srinivasan, M. V., M. Poteser, and K. Kral
1999. Motion detection in insect orientation and navigation. *Vision Research*, 39(16):2749–2766.

Strother, J. A., A. Nern, and M. B. Reiser
2014. Direct observation of ON and OFF pathways in the *Drosophila* visual system. *Current Biology*, 24(9):976–983.

Strother, J. A., S.-T. Wu, E. M. Rogers, J. L. M. Eliason, A. M. Wong, A. Nern, and M. B. Reiser
2018. Behavioral state modulates the ON visual motion pathway of *Drosophila*. *Proceedings of the National Academy of Sciences* , 115(1):E102–E111.

Strother, J. A., S.-T. Wu, A. M. Wong, A. Nern, E. M. Rogers, J. Q. Le, G. M. Rubin, and M. B. Reiser
2017. The emergence of directional selectivity in the visual motion pathway of *Drosophila*. *Neuron*, 94(1):168–182.e10.

Suver, M. P., A. Mamiya, and M. H. Dickinson
2012. Octopamine neurons mediate flight-induced modulation of visual processing in *Drosophila*. *Current Biology*, 22(24):2294–2302.

Suzuki, Y., H. Ikeda, T. Miyamoto, H. Miyakawa, Y. Seki, T. Aonishi, and T. Morimoto
2015. Noise-robust recognition of wide-field motion direction and the underlying neural mechanisms in *Drosophila melanogaster*. *Scientific Reports*, 5:10253.

Sweeney, S. T., K. Broadie, J. Keane, H. Niemann, and C. J. O'Kane
1995. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron*, 14(2):341–351.

Tadmor, Y. and D. J. Tolhurst
1993. Both the phase and the amplitude spectrum may determine the appearance of natural images. *Vision Research*, 33(1):141–145.

Takemura, S.-Y., A. Bharioke, Z. Lu, A. Nern, S. Vitaladevuni, P. K. Rivlin, W. T. Katz, D. J. Olbris, S. M. Plaza, P. Winston, T. Zhao, J. A. Horne, R. D. Fetter, S. Takemura, K. Blazek, L.-A. Chang, O. Ogundeyi, M. A. Saunders, V. Shapiro, C. Sigmund, G. M. Rubin, L. K. Scheffer, I. A. Meinertzhagen, and D. B. Chklovskii
2013. A visual motion detection circuit suggested by *Drosophila* connectomics. *Nature*, 500:175.

Takemura, S.-Y., T. Karuppudurai, C.-Y. Ting, Z. Lu, C.-H. Lee, and I. A. Meinertzhagen
2011. Cholinergic circuits integrate neighboring visual signals in a *Drosophila* motion detection pathway. *Current Biology*, 21(24):2077–2084.

Takemura, S.-Y., A. Nern, D. B. Chklovskii, L. K. Scheffer, G. M. Rubin, and I. A. Meinertzhagen
2017. The comprehensive connectome of a neural substrate for 'ON' motion detection in *Drosophila*. *Elife*, 6:e24394.

Talay, M., E. B. Richman, N. J. Snell, G. G. Hartmann, J. D. Fisher, A. Sorkaç, J. F. Santoyo, C. Chou-Freed, N. Nair, M. Johnson, J. R. Szymanski, and G. Barnea
2017. Transsynaptic mapping of second-order taste neurons in flies by *trans*-Tango. *Neuron*, 96(4):783–795.e4.

Tammero, L. F. and M. H. Dickinson
2002. Collision-avoidance and landing responses are mediated by separate pathways in the fruit fly, *Drosophila melanogaster*. *Journal of Experimental Biology*, 205(Pt 18):2785–2798.

Trischler, C.
2010. Chasing behaviour and optomotor following in free-flying male blowflies: flight performance and interactions of the underlying control systems. *Frontiers in Behavioral Neuroscience*, 4:20.

van der Schaaf, A. and J. H. van Hateren
1996. Modelling the power spectra of natural images: statistics and information. *Vision Research*, 36(17):2759–2770.

van Hateren, J. H.
1992a. Real and optimal neural images in early vision. *Nature*, 360(6399):68–70.

van Hateren, J. H.
1992b. Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *Journal of Comparative Physiology A*, 171(2):157–170.

van Hateren, J. H.
1997. Processing of natural time series of intensities by the visual system of the blowfly. *Vision Research*, 37(23):3407–3416.

van Hateren, J. H. and A. van der Schaaf
1998. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1394):359–366.

van Santen, J. P. H. and G. Sperling
1984. Temporal covariance model of human motion perception. *Journal of the Optical Society of America. A, Optics and Image Science*, 1(5):451–473.

van Santen, J. P. H. and G. Sperling
1985. Elaborated reichardt detectors. *Journal of the Optical Society of America. A, Optics and Image Science*, 2(2):300–321.

Venken, K. J. T., J. H. Simpson, and H. J. Bellen
2011. Genetic manipulation of genes and cells in the nervous system of the fruit fly. *Neuron*, 72(2):202–230.

Vlasits, A. L., R. D. Morrie, A. Tran-Van-Minh, A. Bleckert, C. F. Gainer, D. A. DiGregorio, and M. B. Feller
2016. A role for synaptic input distribution in a dendritic computation of motion direction in the retina. *Neuron*, 89(6):1317–1330.

von Fermi, G. and W. Richardt
1963. Optomotorische Reaktionen der Fliege *Musca domestica*. *Kybernetik*, 2(1):15–28.

von Helmholtz, H.
1867. *Handbuch der physiologischen Optik*, volume 3. Leipzig: L. Voss.

Wainwright, M. J., O. Schwartz, and E. P. Simoncelli
2002. Natural image statistics and divisive normalization: modeling nonlinearity and adaptation in cortical neurons. In *Probabilistic Models of the Brain: Perception and Neural Function*, R. Rao, B. Olshausen, and M. Lewicki, eds., chapter 10, Pp. 203–222. MIT Press.

Weber, F., C. K. Machens, and A. Borst
2010. Spatiotemporal response properties of optic-flow processing neurons. *Neuron*, 67(4):629–642.

Wienecke, C. F. R., J. C. S. Leong, and T. R. Clandinin
2018. Linear summation underlies direction selectivity in *Drosophila*. *Neuron*, 99(4):680–688.e4.

Yamaguchi, S., R. Wolf, C. Desplan, and M. Heisenberg
2008. Motion vision is independent of color in *Drosophila*. *Proceedings of the National Academy of Sciences* , 105(12):4910–4915.

Yang, H. H., F. St-Pierre, X. Sun, X. Ding, M. Z. Lin, and T. R. Clandinin
2016. Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. *Cell*, 166(1):245–257.

Yau, K.-W. and R. C. Hardie
2009. Phototransduction motifs and variations. *Cell*, 139(2):246–264.

Zaghloul, K. A., K. Boahen, and J. B. Demb
2005. Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. *Journal of Neuroscience*, 25(4):860–868.

Zhang, F., L.-P. Wang, M. Brauner, J. F. Liewald, K. Kay, N. Watzke, P. G. Wood, E. Bamberg, G. Nagel, A. Gottschalk, and K. Deisseroth
2007. Multimodal fast optical interrogation of neural circuitry. *Nature*, 446(7136):633–639.

ACKNOWLEDGEMENTS

First and foremost I want to thank Axel Borst for being such a helpful and active supervisor and for letting me be part of his lab. Not only did he manage to reignite the spark of enthusiasm for science in me every time we talked about research projects. He also encouraged me to follow my own ideas, gave me the freedom to pursue side projects and by that offered me the great opportunity to discover and develop my skills and scientific thinking. Always supportive of team work in the lab, he has established an incredibly open and collaborative atmosphere which made my time at the institute such a unique experience for me.

Furthermore, I would like to thank everyone who has participated in my research projects, and all my colleagues, friends and family, which have contributed by accompanying and supporting me throughout this time. Without them, this dissertation would not exist. First, I would like to thank Armin Bahl who was like a second supervisor for me in the beginning and who taught me a very fundamental skill set on how to approach neuroscience, an area that back then was still very new for me. Aljoscha Leonhardt encouraged and initiated many ideas that later developed into fruitful research projects and I am very grateful for countless insightful discussions. Alexander Arenz showed me how true scientific enthusiasm looks like and taught me how to finish a project and write a research article. I want to thank Inês Ribeiro for our collaboration which gave me the opportunity to catch a glimpse into a fascinating and, for me, completely new research question. Many thanks to Florian Richter who not only was on my side during our team projects but also was a very supportive and lovely office mate. Anna Schützenberger for offering her skills in our team project and for being an embodiment of patience and calmness despite all circumstances. A thousand thanks to Nadya Pirogova who did not stop being supportive, helpful, patient and great; without her I would have not made it through some desperate moments of this work. Finally, I want to thank Georg Ammer for teaching me scientific precision. Etienne Serbe and Matthias Meier for being the prime example of teamwork in this lab. Sandra Fendl for patiently explaining me things far beyond my expertise and for countless fun moments. Jesús Pujol-Martí for all the crazy discussions and for sharing many moments with me on the patch rig including the christmas experiment. Tabea Schilling for helping me getting settled in the lab. Stefan Prech for sharing his impressive technical and infinite practical expertise. Jürgen Haag for supporting me with his rich experimental experience.

Last but not least, I want to thank Ruben Portugues and Laura Busse for their advice as part of my thesis committee. Aljoscha Leonhardt, Florian Richter and Nadya Pirogova kindly read and helped me improve this thesis. In the end, I want to thank again the whole Borst department and all of my friendly colleagues around the institute who made the past years so enjoyable, productive, exciting and unforgettable!

AFFIDAVIT

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation **Linear and non-linear receptive fields in *Drosophila* motion vision** 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 **Linear and non-linear receptive fields in *Drosophila* motion vision** is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

17.06.2020	Michael Drews
.....	.....
München, den	Unterschrift
Munich, Date	Signature

CONTRIBUTIONS

MANUSCRIPT I

The temporal tuning of the *Drosophila* motion detectors is determined by the dynamics of their input elements

Alexander Arenz*, **Michael S. Drews***, Florian G. Richter, Georg Ammer & Alexander Borst
*equal contribution

Current Biology, 27(7), 929-944.
doi: 10.1016/j.cub.2017.01.051

A.A., **M.S.D.**, and A.B. conceived the study and designed the experiments. A.A. conducted and analyzed the measurements of T4/T5 cell responses. **M.S.D.** designed the projector-based stimulation arena and performed and analyzed the measurements of the OFF-pathway elements. F.G.R. performed and analyzed the experiments describing the ON-pathway neurons. G.A. performed and analyzed the patch-clamp recordings from lobula plate tangential cells. **M.S.D.** performed the computer simulations. A.A. wrote the manuscript with the help of all authors.

Michael Drews (author)

Alexander Arenz (co-author)

Alexander Borst (supervisor)

MANUSCRIPT II

Visual Projection Neurons Mediating Directed Courtship in *Drosophila*

Inês M.A. Ribeiro, **Michael Drews**, Armin Bahl, Christian Machacek, Alexander Borst & Barry J. Dickson

Cell, 174(3), 607-621
doi: 10.1016/j.cell.2018.06.020

I.M.A.R, B.J.D., and A. Borst conceived and designed this study. **M.D.** performed the functional imaging and data analysis presented in Figure 5. I.M.A.R. performed all other experiments and analyzed the data. A. Bahl and **M.D** .conceived and designed the visual stimuli. C.M. and B.J.D. created MateBook. I.M.A.R and B.J.D. co-wrote this paper with input from all authors

Michael Drews (author)

Alexander Borst (supervisor)

MANUSCRIPT III

Glutamate Signaling in the Fly Visual System

Florian G. Richter, Sandra Fendl, Jürgen Haag, **Michael S. Drews** & Alexander Borst

iScience, 7, 85-95
doi: 10.1016/j.isci.2018.08.019

F.G.R., S.F., and A.B. conceived the study and designed the experiments. F.G.R. conducted and analyzed the imaging experiments for Mi9 and L1. S.F. performed and analyzed all stainings. J.H. performed and analyzed the LPi experiments. **M.S.D.** performed data analysis and model fitting of the receptive fields. F.G.R. wrote the manuscript with the help of all authors.

Michael Drews (author)

Alexander Borst (supervisor)

MANUSCRIPT IV

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 & Alexander Borst
*equal contribution

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 modelling, and wrote the manuscript. All authors participated in editing the manuscript.

Michael Drews (author)

Aljoscha Leonhardt (co-author)

Alexander Borst (supervisor)