UNDERSTANDING SELF-MOTION INDUCED OPTIC FLOW PROCESSING IN LARVAL ZEBRAFISH IN THREE LEVELS

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Yue Zhang
aus Shenzhen / China

Tübingen
2022

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	11.07.2022
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Aristides B. Arrenberg
2. Berichterstatter:	Prof. Dr. Hanspeter A. Mallot

Erklärung / Declaration: Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

"Understanding self-motion induced optic flow processing in larval zebrafish in three levels "

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

I hereby declare that I have produced the work entitled *"**Understanding self-motion induced optic flow processing in larval zebrafish in three levels** "*, submitted for the award of a doctorate, on my own (without external help), have used only the sources and aids indicated and have marked passages included from other works, whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.

Tübingen, den ....

Datum / Date Unterschrift / Signature

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS

3D	three-dimensional
2D	two-dimensional
RF	receptive field
EMD	elementary motion detector
LPTC	lobula plate tangential cells
NMN	neck motor neuron
DNOVS	descending neurons of the ocellar and vertical system
STA	spike-triggered average
SNR	signal-to-noise ratio
MSTd	the dorsal part of the medial superior temporal area
OKR	optokinetic response
OMR	optomotor response
CMN	contiguous motion noise
AMC	anterior medial cluster
GMM	gaussian-mixture model

NCB	nonparametric cluster-based bootstrapping
FoC	focus of contraction
FoE	focus of expansion
EM	electron microscope
dpf	days post-fertilization
APT	area pretectalis
nMLF	the nucleus of the medial longitudinal fasciculus
AF	arborization field

Part I

SUMMARY

1

ZUSAMMENFASSUNG

In dieser Arbeit stelle ich mein Dissertationsprojekt über die neuronale Implementierung der durch Eigenbewegung induzierten Verarbeitung des optischen Flusses im Pretectum des larvalen Zebrafisches. Ich erläutere dies mit David Marrs "Drei-Ebenen"-Framework zum Verständnis des visuellen Systems. Ich gehe die Wichtigkeit dieses Frameworks mit einem Gedankenexperiment an und wende diese anschließend auf das Problem der Verarbeitung des optischen Flusses an. Ich erläutere das rechnerische Ziel, der durch Eigenbewegung induzierten Verarbeitung des optischen Flusses, welches die Extraktion der Eigenbewegungsinformationen aus dem optischen Fluss ist. Um dies zu erreichen, muss der verwendete Algorithmus zwei Berechnungsprinzipien umsetzen: die "*Zerlegbarkeit*" und die "*Identität*". Ich stelle hierfür drei mögliche Algorithmen zur Verarbeitung des optischen Flusses kurz vor und erläutere, wie die beiden Berechnungsprinzipien in diesen Algorithmen umgesetzt werden. Darüber hinaus bespreche ich die relevanten experimentellen Studien und erkläre im Detail, warum die systematische Charakterisierung der rezeptiven Felder der Fliege die neuronale Implementierung des "matched-filter" Algorithmus für die durch Eigenbewegung induzierte Verarbeitung des optischen Flusses aufzeigen kann. Ich erwähne die Vorteile und Grenzen der umgekehrten Korrelation (*reverse correlation*) für die Abschätzung der rezeptiven Felder und stelle eine Methode für die effiziente Schätzung höherer Merkmale der rezeptiven Felder mittels "Merkmalsrauschen" (*feature noise*) vor.

Diese Dissertation beinhaltet drei wissenschaftliche Publikationen. Die wichtigsten Ergebnisse sind in der ersten und dritten Publikation enthalten, während ich in der zweiten Publikation hauptsächlich auf ein spezifisches technisches Problem eingehe. In dem Diskussionsteil diskutiere ich die Vorteile und Probleme im Zusammenhang mit der räumlich-zeitlichen Korrelation des "zusammenhängenden Bewegungsrauschens"(*contiguous motion noise (CMN)*), die in der ersten Publikation entwickelt wurde. Ich erläutere potenzielle Probleme bei der umgekehrten Korrelation und dem zweistufigen nonparametric cluster-based bootstrapping-Test im Hinblick auf die Auswirkungen bei der Interpretation der Ergebnisse der RF-Schätzung. Diese Methoden wurden in der ersten und dritten Publikation verwendet. Außerdem erörtere ich die Einschränkungen der umgekehrten Korrela-

tionstechnik bei der Charakterisierung der funktionalen Eigenschaften, deren Linearität noch zu bestätigen ist.Da die rezeptiven Felder im larvalen Zebrafisch nicht das gesamte Gesichtsfeld abdecken, was eine Abweichung vom ursprünglichen "Matched-Filter-Modell" darstellt, erörtere ich, wie sich diese Abweichung auf die Kodierungsqualität auswirken kann und welche Berechnungsprinzipien dabei eine Rolle spielen. Auch die "Mode-Sensing"-Hypothese und das Pretectum-Hinterhirn-Problem im nachgeschalteten Pfad bei der sensomotorischen Transformation des Sehflusses wird kurz erwähnt. Zum Schluss werden zwei spezifische Lücken auf der Implementierungs-, Algorithmus- und Berechnungsebene diskutiert, die meine Arbeit nicht schließen konnte. Die vorliegende Dissertation gibt nicht nur einen Einblick in die Verarbeitung des optischen Flusses und die Abschätzung der Eigenbewegung bei Zebrafischen, sondern zeigt auch, wie Marrs "Drei-Ebenen"-Framework dazu beitragen kann, unser Verständnis des visuellen Systems auf systematischere und rigorosere Weise voranzubringen.

2

SUMMARY

In this thesis I present my PhD project about the neural implementation of self-motion induced optic flow processing in larval zebrafish pretectum, and dissect it with Marr's three-level framework for understanding visual system. I address the importance of this framework with a thought experiment and applied this framework to the optic flow processing problem. I state one of the computational goal(s) of self-motion induced optic flow processing is to extract the self-motion information from optic flow. The algorithms that can accomplish this goal, must implement two computational principles: the *decomposability* and *identity*. Then I briefly review three putative algorithms proposed for optic flow processing and discussed how the two computational principles are implemented in these algorithms. I also briefly review the relevant experimental studies and explained in detail that why the systematic receptive field (RF) characterization in flies may reveal the neural implementation of the matched filter algorithm for self-motion induced optic flow processing. Then I review the advantages and limitations of the *reverse correlation* method for RF estimation and I introduce the *feature noise* for efficiently estimating the RFs of advanced features.

I include three papers in the summary sections. The major results in my PhD project are in the first and the third paper while I mainly contribute to a specific technical problem in the second paper. In the discussion, I discuss the advantages and problems related to the spatiotemporal correlation of the CMN developed in the first paper. I discuss potential problems in the reverse correlation and the two-step nonparametric cluster-based bootstrapping test used in the first and the third papers in terms of the effects in the interpretation of the RF estimation results. And I also discuss the limitation of reverse correlation technique in characterizing the functional properties whose linearity remains to be confirmed. Since the RFs in larval zebrafish do not cover the entire visual field which is deviated from the original "matched filter model", I discuss how this deviation may affect the encoding quality and which computational principles are involved in. I also briefly mention the "mode-sensing" hypothesis and the pretectum-hindbrain problem in the downstream pathway in the sensorimotor transformation of optic flow. In the end, I discuss two specific gaps across the implementational, algorithmic and computational levels that my work failed to

enclose. I hope this thesis may provide some insight into the optic flow processing and self-motion estimation in zebrafish, but also demonstrate how Marr's three-level framework may help to advance our understanding of visual system in a more systematic and rigorous manner.

Part II

INTRODUCTION

3

INTRODUCTION

We live in a world of images we have of the world, and sometimes it is not the world who is fast changing, but our view of the world. Every active or passive movement of one's body may result in changes of the retinal images of the world. This phenomenon is known as self-motion induced optic flow. Theoretically, this type of optic flow contains sufficient information for determining self-motion status. The neural correlate for self-motion induced optic flow processing have also been discovered in a wide range of animal species. It is very natural to speculate that these neural correlates are involved in the extraction of self-motion information from optic flow. Now we just need to find a way to prove how this is done.

This seemingly easy task involves several difficult questions. First, how can we confirm if an animal uses the self-motion information for any of its behavior at all? And does the neural correlate identified for optic flow processing extract self-motion information instead of other information available in the optic flow? And if and how exactly is the self-motion information represented by the activities of the neural correlates in an unambiguous, decodable way? I consider these questions difficult not because they require advanced experiment setups or fancy data analysis methods, but because they are addressing different aspects of a complex system at the same time and thus required to be "decomposed" before answering.

Similar questions have been discussed in the well-known article "could a neuroscientist understand a microprocessor" by Jonas and Kording[38], in which they applied some common experimental and data analysis methods in neuroscience to study the function of a microchip. The results they obtained are all somewhat related with some computational properties of the components in the microprocessor, but none of them can reveal the goal of the computational process executing in the microprocessor. They believed it is mainly due to the inappropriate analysis methods. However this is not entirely the case, at least not for the problems I met in my PhD study. These problems are mostly due to the mismatch between the level of the complex system the questions are addressing, and the level addressed by the observation I made. The "level" I am referring to here is a concept introduced by the meta-scientific framework first pro-

posed by Werner Reichardt and Tomaso Poggio, and finalized by David Marr[60,61,73]. This framework suggests visual system, as a complex system consists of three loosely related levels, and one can only fully understand a visual process by linking the observation acquired at different levels together. And the problems I met were untangled and became answerable by applying this three-level framework.

In view of these, I would like to present my PhD project as a demonstration of Marr's three-level framework in this thesis. Before discussing why a systematic characterization of motion sensitive neuron receptive field may help to address the optic flow problems mentioned above with Marr's three-level framework, I will start the introduction with a thought experiment demonstrating what is the "difficult question" I mentioned above in a more explicit way.

3.1 A PARALYZED ALIEN AND THREE LEVELS OF UNDERSTANDING

> "A wing would be a most mystifying structure if one did not know that birds flew. "
>
> - Horace Barlow, 1961[2]

the "paralyzed alien" thought experiment

A group of world-class neuroscientists were led into a room where they found a cage covered with black cloth. The man in black standing next to the cage said to the crowd: "You are invited to here because we found this extra-terrestrial animal in the UFO we shot down yesterday." He uncovered the cage and there is an unknown creature lying in the cage. "It survived from the spaceship crash, but it is completely paralyzed. We cannot determine what caused the paralysis and treat it because we cannot take off its exoskeleton suit which also blocks all electromagnetic waves for any medical imaging. So in short, it is permanently paralyzed and there is nothing we can do about it."

"The good news is its brain is not covered by the suit, which gave us a chance to study it with these two powerful setups..." The man in black pointed at a fancy setup on his left, "This setup can provide all kind of physical stimuli. You can use it to present any sensory stimuli or directly stimulate the brain in any way at any resolution and scales. And this-" He pointed at the setup on the right "-allows you to record any physiological signal and anatomical structure from its brain, again, at any resolution and scales. You can forget about all technical limitations from now

on." "And what I want you to determine is considerably simple" He walked back to the cage, "can this paralyzed alien see at all?" Let's assume these neuroscientists can solve any problems instantly and never make a mistake. Can they answer this question when the alien stays paralyzed?

An empirical definition of "seeing" in neuroscience

When I discussed this thought experiment with my colleagues, they immediately pointed out the lack of definition of *"seeing"* in this thought experiment. What does "seeing" mean to us? If some parts of the alien are responding to visual stimuli, is it sufficient to conclude the alien can see? To extend this further, can we know anything about the alien's perception without behavioral evidences? These questions matter not only in a metaphysical way, but also in a very practical sense: before every experiment, I need to screen out the fish with no vision problems for functional characterization of visual systems. Examine the structural integrity of visual system provides a way for screening, however it cannot determine the functional status of the visual system. A blind fish may have an structurally intact but malfunctioning visual system. To solve this problem, most if not all functional visual neuroscience studies adopt a behaviorism solution, which screens animal by their behavioral responses to visual stimulus, such as the optokinetic response (OKR) and optomotor response (OMR)[7,65]. The animals repeatedly responding to visual stimuli with the expected behaviors are believed to be able to see.

In my opinion, without behavioral evidence support, any assertion about animal perception may be controversial. For example, in the "paralyzed alien" thought experiment, if we conclude the alien is able to see based on some neuron-visual feature correlations, we should also conclude that the computer detecting the same features can see as well[1]. This does not mean all reductionism investigations are pointless. Quite the opposite, they may deepen our understanding of the computational mechanism of the sensory systems, and allow us to predict some computational features once the computational goal is known.

Similarly, revealing the computational goal of a visual process may not help to determine the function of individual components either. These imply some necessary "compartmentalization" in understanding visual systems: the observation in one "compartment" cannot directly lead to conclusions in other compartments. This has already been well stated in the quote from

1 As Poggio and Marr concluded: "*The underlying philosophical issue here is that reductionism does not imply constructionism*"[60].

Computational	The goal and the strategy of computation
Algorithmic	The algorithm/representation employed by the system
Implementational	The hardware implementation of the algorithm/representation

Table 3.1: Marr's three-level framework

Poggio and Reichardt's three levels of understanding framework

Horace Barlow above. Werner Reichardt and Tomaso Poggio addressed this compartmentalization as different levels for understanding complex systems[73]. Despite "... the knowledge at one level might be of little help at any other one. However, an understanding of the functional property at higher levels might help to develop approaches at lower levels."[73]. It seems this compartmentalization is more of a feature of our cognition instead of a property of the sensory system. I would not further expand this topic in this neuroscience thesis as it is more appropriate to discuss under philosophical, psychological or semiotics frameworks, instead, I will focus on how this compartmentalization may help for rigorous investigation of visual system.

Visual system or any other complex systems, in Weiner and Poggio's opinions, should be understood at three levels which concern the function of the entire visual system, the subsystems and the neural circuits or individual neurons respectively. These three levels are divided based on the investigation scale which may not well reflect the compartmentalization in our cognition process of visual system. Therefore it has very limited power in explaining the problem in the "paralyzed alien" example, while Marr's framework may be more suitable to unravel the confusions in the thought experiment. The definition of the three levels of understanding visual systems in Marr's framework is summarized in Table 3.1.

Marr's three-level framework

In plain language, the computational level is asking "*what does the system do?*" The algorithmic level asks "*How does the system work?*" And the implementational level concerns "*Why the system can work with its current hardware configuration?*" Marr's three-level framework corresponds better to our cognitive process if we recall the first time we saw chopsticks: the descriptions "they are wooden sticks" or "they can be held and move with your fingers" will not help to understand their function or their *computational goal*. Once the function is known, the next question will usually concern the *algorithm* (how to use them to deliver food). And after these are known, its hardware implementation may

start to make sense (e.g. the shape and the material of the sticks).

Marr stated, as we see in the example above, that the computational, algorithmic and implementational level are loosely related. Although the conclusion at one level may not directly determine property at another level, it may still leave us some hints. This loose relationship sometimes tangles the three level up. To avoid this in my thesis, I will introduce the optic flow processing by first defining its computational goal and principles before I introduce the putative algorithms and their neural implementations.

3.2 OPTIC FLOW PROCESSING IN THREE LEVELS

Optic flow is generally defined as the movement of luminance patterns in a two-dimensional (2D) image[5,22,42,43] projected on a three-dimensional (3D) surface. It is an objective phenomenon which should not be confused with *visual motion*, the perception yield from optic flow processing. It can be represented by as a vector field where each vector represents the movement velocity of local image patterns (Figure 3.1). Optic flow can be categorized by its source, scale and the visual perception triggered. For example, they are resulted from the real object movement, or apparent motion. The former type of optic flow can be further divided into the optic flow induced by the movement of the external objects and the ones induced by the movement of the observer. The terms of different categories of optic flow grouped by the scale or visual perception triggered are often interchangeably used which may result some confusion. For example, the wide-field optic flow is sometimes considered to be equivalent to the optic flow triggering global motion perception. This is not necessarily the case as there are evidences suggesting the perception of global motion may also be triggered by the optic flow presented in local regions[41]. And the incoherent wide-field optic flow may not lead to global motion perception[1,14]

the definition of optic flow

To understand optic flow processing with Marr's three-level framework, we need to answer the following questions:

1. What information in what type of optic flow is used by animals, and in which behaviors? *(computational goal)*
2. What structure/form/properties must the information be extracted from optic flow for the computational goal? *(computational principles)*
3. How can this extraction and structuring be done with finite number of steps and computational components?

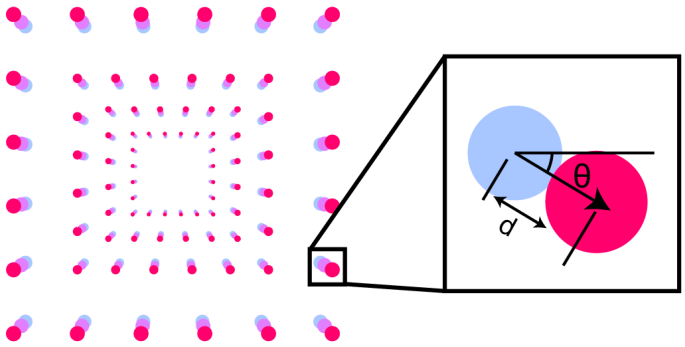


Figure 3.1: The optic flow estimated as the displacement field of two images indicated by the velocity vectors. The d and θ indicate the norm and the phase angle of the selected vector

How efficient and robust is the corresponding algorithm in extracting and preserving the information from optic flow? *(algorithm)*

4. How is the algorithm for optic flow processing implemented in what functional structure (i.e. receptive fields) of which neurons? How may the limitations in the signal quality, computational efficiency of the neural implementation affect the robustness and efficiency of the algorithm? *(neural implementation)*

In this thesis I will discuss these questions in relation with the optic flow induced by self-motion, which is largely overlapped with the optic flow triggers the perception of self-motion. Gibson was the first to point out it is sufficient to determine self-position, heading direction and speed of the observer from this type of optic flow[2][21,22]. This is confirmed by the later studies which have demonstrated multiple ways to extract this information from global optic flow fields[10,11,42,43,56,71]. It is natural to speculate that at least one of the computational goal of self-motion induced optic flow is to infer the 3D self-motion status from the 2D optic flow field.

The computational goal of self-motion induced optic flow

This speculation has been proved in human[39,93,94,106]. However the investigation strategy in most of these studies cannot be directly transferred to the other species which we cannot easily communicate with. For the other animals, we have to assess this hypothesis by finding the behavioral task that cannot be accomplished without knowing self-motion status. In most species, it is evaluated with the visual stabilization behavior such as gaze, head orientation, body posture and body position stabilization[3,35,49,76]. These behaviors all share a common

2 The optic flow induced by self-motion is not equivalent to the self-motion triggered optic flow as the latter can also be apparent motion.

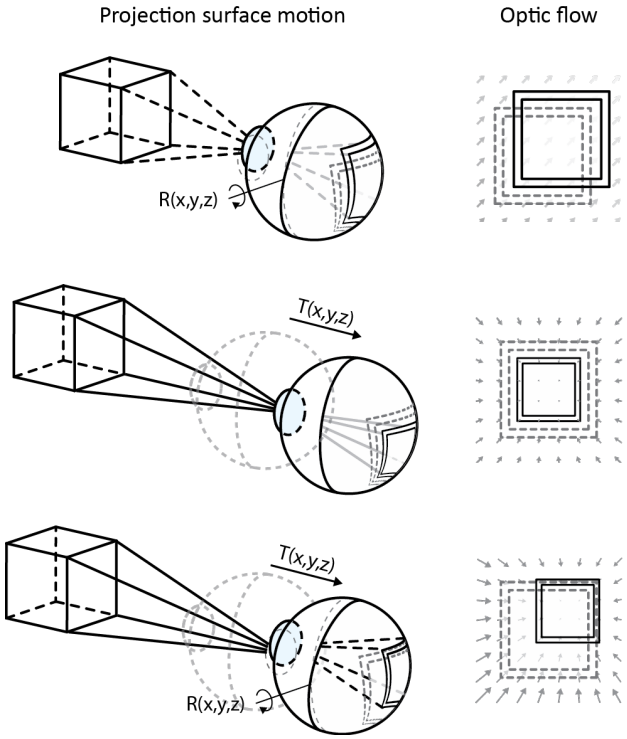


Figure 3.2: Three types of self-motions (left column) and the retinal projection of the optic flow they induced (right column). $T(x, y, z)$ and $R(x, y, z)$ indicate the velocity vectors for self-translation and self-rotation respectively.

goal which is to restore the previous self-motion status from the perturbation. And they can be robustly triggered in most animals studied without visual problems. Ideally the speculated computational goal may be confirmed in the animals able to perform these stabilization behaviors when the optic flow is the only available sensory cue. However it is hard to block the other sensory inputs, especially when the animal is moving. Therefore most of these studies "fake" the perturbation by presenting wide-field apparent motion whose vector field is identical to the optic flow field resulted by a perturbation. If animals' behavioral responses to the apparent motion may restore their previous self-motion state from the corresponding illusionary perturbation, then we can prove self-motion information has been inferred from optic flow by the animals tested. This apparent motion induced self-stabilization behavior has been observed in flying insects, fish, birds and monkeys[3,23,35,49,54], which suggest the inference of self-motion status is a common computational goal of optic flow processing across these species.

A rigid object can only be moved in one translation direction and rotate around one axis at one time and the information of self-translation and self-rotation cannot be mixed. If an object is translating and rotating at the same time, the optic flow field

The computational principles of self-motion induced optic flow processing

resulted will be the resultant vector fields of the optic flow fields induced by the translation and rotation separately (Figure 3.2). Based on these facts, I assert two computational principles for self-motion induced optic flow processing:

1. *Decomposability*: The translational and rotational components in self-motion induced optic flow must be processed or represented separately.
2. *Identity*: Every velocity vector in an instantaneous self-motion induced optic flow field may correspond to one and only one translational and rotational self-motion vector.

If the *decomposability* principle is not satisfied, the information of self-translation and self-rotation extracted from optic flow will be confused, and animals may fail to call the correct motor program for stabilizing their self-motion status by compensating the rotation and translation components in perturbations. If the *identity* principle is not satisfied, animal may fail to determine the velocity of translation and rotation to be compensated for self-stabilization. However, satisfying these two principles does not always mean the system can recover its translation and rotation status. In Figure 3.3, all self-motion induced optic flow fields are mapped to a 2D configuration space by the weight of the translation and rotation components in these flow fields. The optic flow processing systems are simplified into a process which converts optic flow to representations of translational and rotational self-motion. The four systems in A satisfy the *decomposability* principle while the ones in B do not. Except the bottom right system, the other systems in A will fail to represent the self-translation and rotation simultaneously. The problem here is not at the computational level, but the algorithmic and implementation level.

the analytical model of self-motion induced optic flow processing

Before I start to discuss the optic flow processing at the algorithmic level, I shall quickly introduce the analytical model developed by Koenderink and van Doorn[42] for describing the translation- and rotation-induced optic flow. Let $\vec{t}$ and $\vec{r}$ be the velocity vectors for self-translation and rotation motion. For a set of visual cues (or fiducial points) at the position $p_1, p_2, ...p_n$, the corresponding optic flow vector $\vec{T}$ and $\vec{R}$ at position p_i is calculated as:

$$\vec{R} = -\vec{r} \times \hat{p}_i \tag{3.1}$$

$$\vec{T} = -\frac{\vec{t} - (\vec{t} \cdot \hat{p}_i)\hat{p}_i}{|p_i|} \tag{3.2}$$

where the $\times$ and $\cdot$ indicate the cross and the dot products of vectors. The $\hat{p}$ is the normalized unit vector of p by dividing

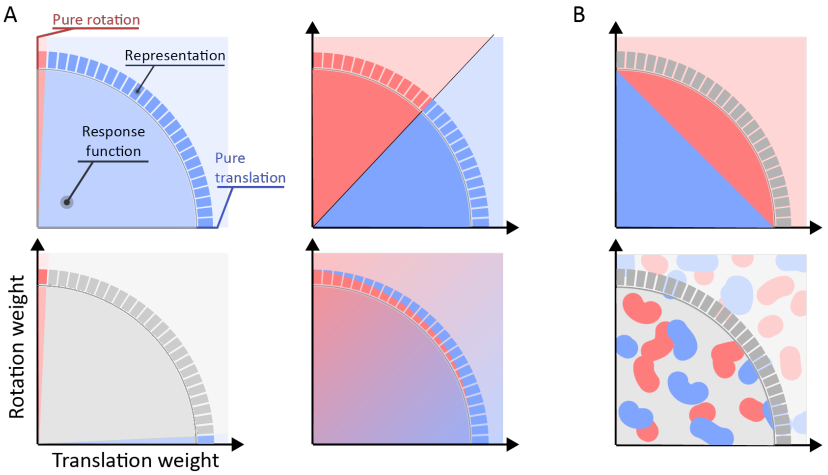


Figure 3.3: The example systems who satisfy (A) or fail to satistify (B) the principle of decomposability are presented in a configuration space composed of translation and rotation weights. The weights represent the significance of the corresponding component in the optic flow field. The optic flow processing systems is simplified as a response function to the translation- and/or the rotation-induced component. The blue and red color indicates the responsiveness to the translational and rotational components respectively. The rectangles in the middles are the representations of self-translation and self-rotation status formed by the corresponding systems. For example, the system represented by the top-left plot in A only respond to the rotation-induced component when the translational component is absent.

with its norm $|p|$. As mentioned before, when the observer is moving and rotating at the same time, the self-motion induced optic flow $\vec{M}$ is simply the sum of the optic flow field induced by the translation movement $\vec{T}$ and the rotation movement $\vec{R}$:

$$\vec{M} = \vec{R} + \vec{T} \tag{3.3}$$

$$= -(\vec{r} \times \hat{p}_i) - (\frac{\vec{t} - (\vec{t} \cdot \hat{p}_i)\hat{p}_i}{|p_i|}) \tag{3.4}$$

Putative algorithm for optic flow processing

It is clear in equation 3.4 that under the constraints set by the two computational principles, the self-motion induced optic flow processing problem in principle is a polynomial regression problem which should be quite easy to solve. Several algorithms have already been proposed in the last century. For example, the analytical solution Longuet-Higgins and Prazdny[56] proposed can determine self-translation and rotation status from the velocity difference of two visual cues in the same area of the visual field but at different depths. The algorithms developed by Koenderink and van Doorn[42] determines self-motion status in an iterative way. In each iteration, the algorithm computes the difference between the observed optic flow field and the flow field yield from the estimated self-translation and rotation vector. And based on this difference, this recursive algorithm may find the least squared estimation of self-motion status. Dahmen, Franz and Krapp[10,11,18] proposed the "matched filter" algorithm which transforms the optic flow field into the scores of a set of linear filter that are highly similar to certain translation- or rotation-induced optic flow fields ($\vec{F^T}$ and $\vec{F^R}$ respectively). For the optic flow field $\vec{M}$ with n elements, the transformed score s is calculated as:

$$s = \vec{M} \cdot \vec{F} \tag{3.5}$$

$$= \sum_i^n \vec{m}_i \cdot \vec{f}_i \tag{3.6}$$

$$= \sum_i^n -(\vec{r} \times \hat{p}_i + \frac{\vec{t} - (\vec{t} \cdot \hat{p}_i)\hat{p}_i}{|p_i|}) \cdot \vec{f}_i \tag{3.7}$$

The implementation of decomposability and identity in the matched filter algorithm

If the filter $\vec{F}$ matches the optic flow induced by rotation $\vec{r^*}$ that $\vec{f}_i = \vec{r^*} \times \hat{p}_i$, and all visual cue are spherically symmetrically distributed with the same depth, let $\vec{r}, \vec{r^*}$ and $\hat{p}_i$ be unit vectors[3]:

$$s = \sum_i^n -(\vec{r} \times \hat{p}_i + \frac{\vec{t} - (\vec{t} \cdot \hat{p}_i)\hat{p}_i}{|p_i|}) \cdot (\vec{r^*} \times \hat{p}_i) \tag{3.8}$$

$$= \frac{2n(r \cdot r^*)}{3} \tag{3.9}$$

3 This constraint is only for the simplicity but not necessary, if they are not unit vector, the function will still be linear that $s = a(r \cdot r^*) + k$

Similarly for the filter matching the optic flow field induced by the unit translation vector $\vec{t^*}$, we can write a similar equation:

$$s = \sum_i^n -(\vec{r} \times \hat{p}_i + \frac{\vec{t} - (\vec{t} \cdot \hat{p}_i)\hat{p}_i}{|p_i|}) \cdot (\frac{\vec{t^*} - (\vec{t^*} \cdot \hat{p}_i)\hat{p}_i}{|p_i|}) \quad (3.10)$$

$$= \frac{2n(t \cdot t^*)}{3} \quad (3.11)$$

Equation 3.9 and 3.11 show the matched filter transforms self-translation/-rotation induced optic flow components into the inner products of the self-motion vectors inducing the optic flow and the self-motion the matched filter represents. Theoretically the self-motion status may be fully restored with at least 3 translational and 3 rotational linear filters. These equations also demonstrated that the decomposability and identity principles are satisfied in the matched filter algorithm. Again this does not mean they can always recover the full self-motion status from global optic flow. The example systems shown in Figure 3.3 might seem poorly designed as they fail to persevere all self-motion information, however similar problems may also be found in these three delicate algorithms, especially when one start to consider the possible neural implementation for these three algorithms.

The detection problems in the neural implementation of optic flow processing algorithms

First, the motion vectors in the optic flow fields need to be inferred from the continuous visual input. The motion inference precision is critical for the estimation accuracy of the algorithm proposed by Longuet-Higgins and Prazdny[56]. A slight error in the speed or direction estimation of the local velocity vectors may result in a significant error in the estimation of self-motion status. The second algorithm has a different problem: it is relatively easy to implement in a recursive neural circuit, however the iterative computation is very time-consuming, especially when the signal-to-noise ratio (SNR) of neural signal is poor. In the extreme cases, it is questionable whether this algorithm will ever converge.

The matched filter algorithm is more robust to the errors in the visual motion detection as it integrates motion from the entire visual field for self-motion estimation. It is also relatively efficient because of the low computational complexity. However it has very little power in estimating the translation or rotation speed. It requires an additional nonlinear speed tuning algorithm to estimate this information. This is also indicated in the equation 3.9 and 3.11 that without the nonlinear speed tuning algorithm, the self-motion estimated with the matched filter algorithm will be ambiguous in terms of the speed and direction

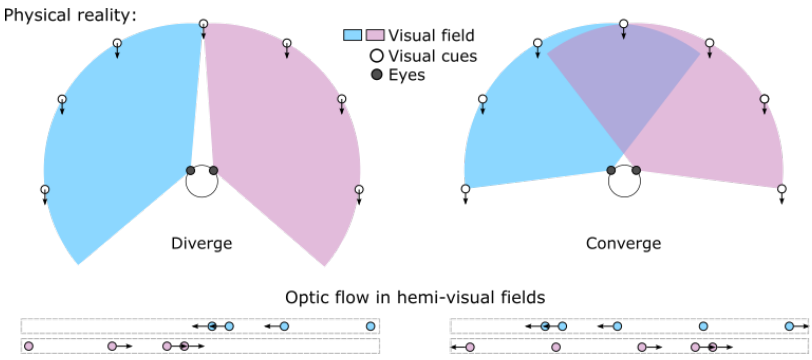


Figure 3.4: The lower plots show the optic flow induced by the same forward-translation (illustrated by the top plots) projected on the simplified one-dimensional left and right retina. Note the difference induced by the eye convergence in the right column.

information. Also the conditions I set for equation 3.9 and 3.11 may be an oversimplification for the optic flow processing in naturalistic scenes. Most times the visual cues are rarely symmetrically distributed and equidistant from the observer, which might also increase the ambiguity of the self-motion information estimated.

In addition, all these three algorithms may not solve the problem caused by the vergence eye movements illustrated in Figure 3.4. Most animals investigated in optic flow research, except some flying insects whose eyes are fixed, are able to perform vergence eye movements[53]. The composition of optic flow field changes during this eye movements. For example, the focus of contraction (FoC) in the forward-translation-induced optic flow in Figure 3.4 become an expansion point. These problems only start to appear at the implementation level, and many of them may greatly reshape our understanding of the algorithms.

These are all minor issues in comparison with the last problem that all these algorithms are completely hypothetical, which will not help us to understand optic flow processing at all until the corresponding neural implementation is found. Only when this correspondence has been revealed, the understanding at the algorithmic level will be a tremendous help to develop a viable understanding of the neural implementation. In the next section I will briefly review the important findings of the neural implementation related with optic flow processing in previous studies.

3.3 VISUAL PROCESSING OF GLOBAL OPTIC FLOW

The neural correlates of self-motion induced optic flow have been found in the visual system of most animals with visually guided self-stabilization behaviors, for example, the lobula plate tangential cells (LPTC) neurons in flies[35,46,88], the neurons in the lobula and medulla in the optic lobe of the shore crabs[32], the pretectal neurons in zebrafish[44,49], the neurons in the pretectal nucleus lentiformis mesencephali of pigeon[99–101], and the the dorsal part of the medial superior temporal area (MSTd) area of the visual cortex in primates and human[63,85,87,93,94]. In fly, shore crab and primates, it is found the responses of these neural correlate is mainly modulated by the self-motion related properties in the optic flow, but weakly influenced by the static features such as the visual cue texture[32,87]. Most of these neurons can be classified by their preference to the type of self-motion induced optic flow (translation induced vs. rotation induced)[40,46,49,85,99]. However there are some exceptions: in pigeon, rabbit and primate, the rotation sensitive neurons were found also sensitive to the shearing motion[82,87,97]. And in primate the preferred translation direction of some translation sensitive neurons are also modulated by the rotation-induced optic flow[85].

The ratio of translation or rotation selective neurons varies a lot across species: The optic flow sensitive neurons found in zebrafish larvae and shore crabs are predominantly translation sensitive neurons[32,49]. In flies it is the opposite: there are much more the rotation sensitive neurons than the translation sensitive neurons[35,46]. In pigeon, primate and human the number of translation and rotation sensitive neurons are similar[63,85,99]. The ratio of translation/rotation selective neurons and their tuning profiles may relate to the computational-level questions mentioned in the previous section. And the variation of these properties among different species could imply differences in the neural implementation or the algorithm selected for optic flow processing. However they might also be explained by the differences in visual statistic or the differences of the motor programs involved in self-stabilization behaviors. Thus these characterizations and variation may not be used as determinative evidences for revealing the optic flow processing algorithm.

But what should be considered as useful evidence? If we take the reductionism view of the brain that each individual neuron serves as a computational component of one or multiple algorithms, we will get a reductionism answer for this question: systematic characterization of the responses function of indi-

vidual neurons in the neural correlate of optic flow processing. This may not be the best answer for experimental neuroscientist as this type of experiments is often very hard to perform in many cases. In some vertebrate species it is not even possible to record the entire neural correlate at single neuron resolution. It is slightly easier to perform these experiments in flying insects because of their relatively small neural population size for optic flow processing. This makes fly the first well studied animal model of optic flow processing at the algorithmic and implementational level.

Optic flow processing pathways in flies

As I mentioned in the previous section, the robustness and efficiency of many optic flow algorithms are more or less depending on the inference precision of local visual motion. In flies, the motion information is gradually extracted by the neurons in lamina, medulla and lobula from the raw image sequence[15,37,74,86]. These neurons are believed to serve as the Reichardt motion detector and in literatures they are often addressed as elementary motion detector (EMD)[4,17,27,72,73,105]. The Reichardt detector is an algorithm for robust motion detection from the spatiotemporal correlation of image sequences[25]. But its power is limited to the detection of first-order motion, and it is tuned to the temporal frequency instead of the absolute speed of visual motion[25,103]. Therefore the signal emitted from these EMDs cannot directly represent local velocity but requires to be further integrated in the downstream neurons. A few decades ago, the LPTCs were found tuned to self-motion induced optic flow instead of local visual motion[24,29]. The LPTCs project to neck motor neuron (NMN) via descending neurons of the ocellar and vertical system (DNOVS)[26,35,95] which is involved in the gaze stabilization behavior[35], suggesting the LPTCs play a role in the computational process whose goal involves self-motion extraction. This behavioral relevance is further supported by the disruption of self-stabilization behaviors after the removal of LPTCs[20,28,30]. These evidences indicate LPTCs may be the pivot for self-motion induced optic flow processing.

optic flow processing algorithm revealed by systematic receptive field characterization in flies

In 1996, Holger Krapp and Roland Hengstenberg characterized the RF of these LPTCs in blowfly as a first step for comprehensive characterization of LPTC's stimulus-response function[46]. I will expand to talk about the concept "receptive field" in the next section. Here it can be roughly considered as the linear part of the stimulus-response function whose structure may reveal the spatiotemporal integration of local motion information. In this paper, the RF is measured by the single electrode recording of membrane potential responses to local motion stimulus in

each individual neuron[46,47]. The RF structure is visualized as a quiver plot, in which the preferred direction in each local region mapped is represented by the phase angle of vector in the corresponding region. It was found the RF structures of these optic flow sensitive neurons are very similar to certain translation- and rotation-induced optic flow, as the matched filter algorithm predicts[10,18,45,46]. Furthermore, these RFs measured by their responses to local visual motion stimuli[47] can predict the neural responses to the self-motion induced optic flow stimuli at the global scale to a high extent[40]. Similar RF structure is also found in the downstream premotor and motor neurons involved in gaze stabilization neurons, which are presumably inherited from the LPTCs[35,95]. Together these evidences suggest that the algorithm fly employed for self-motion induced optic flow processing is very similar to the matched filter algorithm.

This investigation strategy however cannot be directly transferred to the vertebrate species with much more neurons involved, as it is almost impossible to record the membrane potential for every single neuron in these species. For these animals, we can only study the neural implementation of the optic flow processing algorithm with large-scale neurophysiology recording method such as calcium imaging. In comparison with the electrophysiology recordings, calcium imaging data has lower SNR and slower kinetic. And it cannot record the subthreshold activities which is required in the RF estimation method in fly because local stimulus might often be insufficient to trigger suprathreshold activities, especially in the optic flow sensitive neurons with large RFs. If we want to continue the investigation of the optic flow processing algorithm with the RF estimation approach, it is necessary to improve the RF estimation method to be compatible with *in vivo* calcium imaging. In the next section I will discuss how reverse correlation, one of the most common RF estimation techniques can be extended for efficient motion RF estimation.

3.4 REVERSE CORRELATION AND ADVANCED NOISE STIMULUS

The concept receptive field is originally defined as a physical region (e.g. a skin area or a region in the visual field) involving in the sensory stimulus modulation of certain neuron activities[33,81]. It has been generalized to describe correlations between certain neural activity events and sensory stimulus. Given the high di-

mensionality of the possible features to be included in a sensory stimulus, it is almost impossible to have a comprehensive RF characterization of the neuron of interest without knowing its precise function. We must narrow down the scope of features to be estimated with our prior knowledge of neurons' functions. And the RF estimated based on these prior knowledge should be only considered as the approximation of the "ground-truth" RF[4]. Therefore the goal of receptive field estimation should not be to recover the complete stimulus-response function but to evaluate the response function in a defined computational task.

The problems in "comprehensive RF estimation"

Even under this constrain, it may still take too much time to estimate RFs in a brutal way: if a neuron is responding to stimuli contains n features or parameters, and each parameter can take m values or states, then there are m^n possible stimuli to be presented to the neuron for RF estimation. Each of these possible stimuli can be represented by a high-dimensional vector in a configuration space of stimulus parameters. The partition at each dimension of this stimulus vector is defined by the value of the stimulus parameter the dimension corresponds to. If a neuron responds to similar stimuli in a similar way, then its RF may be well estimated by a subset of visual stimuli that are uniformly distributed and cover the entire response-triggering field in the stimulus parameter space of the neuron. Since the relationship between stimulus parameters which is represented by the phase angle of stimulus vectors, instead of their absolute value is more important for interpreting RF estimations, the constraint of uniformity can be further relaxed to the uniform phase angle distribution of the stimulus vectors with same length (examples shown in Figure 3.5).

This strategy is similar to what is employed by the reverse correlation method. The reverse correlation technique, sometimes known as spike-triggered average (STA) has been widely applied to estimate the RF of neurons believed to encode simple sensory features such as edge orientation in visual stimuli and temporal frequency in auditory stimuli[90,104]. The reverse correlation requires the stimulus parameter space is an Euclidean space, that the stimulus features are independent, measurable parameters[9]. The receptive field f of a neuron can be estimated from a set of independent stimulus evoked responses by:

Reverse correlation as a linear method for RF estimation

4 We cannot know if our prior knowledge about the function of a neuron is complete without a comprehensive functional characterization which is impossible to perform. If it isn't complete, then the RF estimated will be biased by the prior.

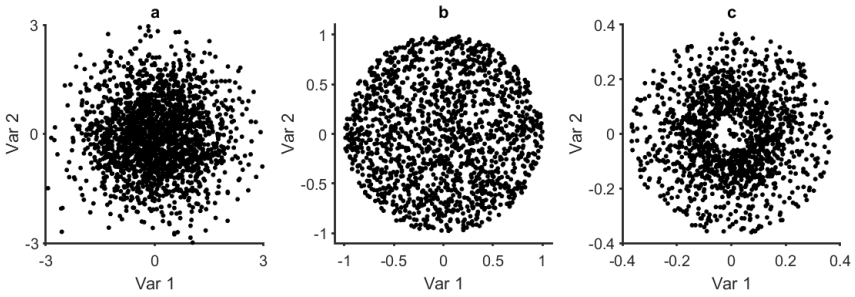


Figure 3.5: Scatter plots of samples from three spherically symmetric probability distribution: **a**: a 2D Gaussian distribution; **b**: a circular uniform distribution with uniform radial distribution; **c**: a circular uniform with Rayleigh radial distribution. Each dot represents a 2-element stimulus vectors.

$$\hat{f} = -\frac{1}{n}\sum_{i=1}^{n} r_i s_i \tag{3.12}$$

$$n = \sum(r_i) \tag{3.13}$$

r_i is the neural responses (e.g. number of spikes) during the displaying period of the i^{th} stimulus (s_i) that should be a non-negative real number. The estimation $\hat{f}$ is the average pattern of all stimulus patterns that triggered neuron responses.

Two assumptions for unbiased RF estimation with reverse correlation

Besides the general independency assumption of stimulus parameters and neural responses, there are another important assumptions that need to be satisfied for unbiased reverser correlation estimation:

1. The stimulus set is spherically symmetric.
2. The nonlinear transformed neuronal response ($NL(r \cdot f)$) is positively proportional to the dot product values ($r \cdot f$) that are greater than 0 as shown below.

$$NL(x) = \begin{matrix} ax, & \text{x} > 0 \\ 0, & \text{else} \end{matrix} \tag{3.14}$$

The first assumption is similar to the uniformity constrain of the stimulus phase angle distribution mentioned above. Chichilnisky[9] explained the spherical symmetry assumption as follows: if the stimulus set is spherically symmetric, for any stimulus vector s_i, it is always possible to find a vector s_i^* with the same length and symmetric in related to the RF vector f. The sum of s_i^*

and s_i will thus be multiples of f (kf). The reverse correlation estimation of the RF can be therefore written as[5]:

$$\hat{f} = \frac{1}{n}\sum_{i=1}^{\frac{n}{2}} s_i r_i P(s_i) + \frac{1}{n}\sum_{i=1}^{\frac{n}{2}} s_i^* r_i^* P(s_i^*) \tag{3.15}$$

Given that the probability distribution of the stimulus set ($P(s)$) is spherical symmetric, we have:

$$\hat{f} = \frac{1}{n}\sum_{i=1}^{\frac{n}{2}} s_i r_i P(s_i) + \frac{1}{n}\sum_{i=1}^{\frac{n}{2}} s_i^* r_i^* P(s_i^*) \tag{3.16}$$

$$= \frac{1}{n}\sum_{i=1}^{\frac{n}{2}} (s_i + s_i^*) r_i P(s_i) \tag{3.17}$$

$$= \frac{1}{n}\sum_{i=1}^{\frac{n}{2}} (s_i + s_i^*) NL(s_i \cdot f) P(s_i) \tag{3.18}$$

$$= f\frac{1}{n}\sum_{i=1}^{\frac{n}{2}} kNL(s_i \cdot f) P(s_i) \tag{3.19}$$

When the second assumption is satisfied, $NL(s_i \cdot f)$ will be zero when $k < 0$. And because the stimulus is spherical symmetric, we can always find another pair of stimulus $\mathbf{s}$ and $\mathbf{s}^*$ that are symmetric to the hyperplane perpendicular to f. The sum of this pair of vectors will thus be $-kf$. Therefore the $\frac{1}{n}\sum_{i=1}^{\frac{n}{2}} kNL(s_i \cdot f)P(s_i)$ will be a positive real number, and thus the reverse correlation estimation $\hat{f}$ is positively proportional to f. If the spherical symmetry assumption is not met, the estimation will be biased as demonstrated in Figure 3.6.

Some advanced visual features may "emerge" from luminance noise stimulus, for example, "edge" is a feature which requires the detection of luminance and contrast in local area. And visual motion requires the detection of the spatiotemporal correlation in the luminance pattern sequence. A problem is that, even in a spherical symmetric luminance noise stimulus, the distribution of these advanced features are often not spherically symmetric (Figure 3.7). Another problem is the frequency of these advanced features are much lower than the basic features in the noise stimulus, which result in the under-representation of these advanced features in the RF estimated.

One can perform whitening transformation to correct these biases[6], however, this transformation usually comes with the

5 The equations 3.14 to 3.19 are adapted from [9]

6 Technically it is not as easy as it sounds. The stimuli of basic feature must be first converted to the stimuli of advanced feature. This process may involve more assumption of the data statistics to be satisfied

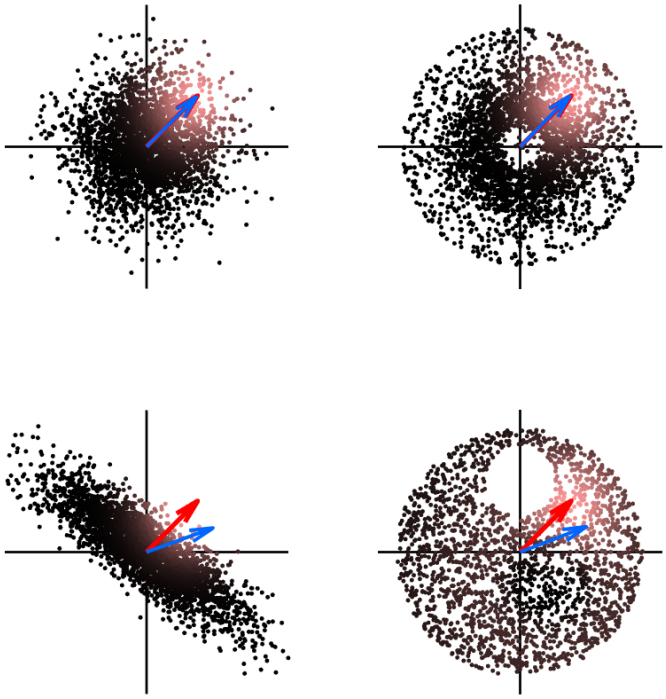


Figure 3.6: The ground truth RF vector (red) and the RF vector estimated with reverse correlation (blue) in the stimulus parameter space. The noise stimuli in the top row are spherical symmetry and the ones in the bottom are not spherical symmetry.

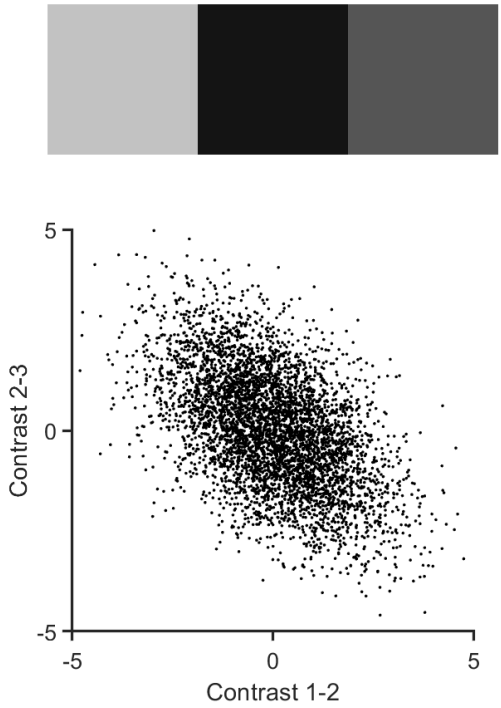


Figure 3.7: Top: An instance of a 3-pixel noise stimulus, the pixel brightness is spherically symmetric. Bottom: Each dot indicates the contrast (brightness difference) of the neighbor pixels.

The advantages of using advanced feature noise stimulus in advanced feature RF estimation

cost of estimation quality. Another more straightforward way is to directly construct a stimulus that is spherically symmetric in the parameter space defined by the advanced feature of interest but not the basic visual features. In other words, this method shifts the estimation bias from the feature of interest to other features. The benefit of replacing the basic sensory feature with the advanced feature of interest is the lift of estimation efficiency especially when the SNR is poor. If a neuron is responding only to a certain advanced feature emerged from the basic feature in the sensory stimulus, it will respond more frequently to the advanced feature noise than the basic feature noise, which in turn leads to a higher information acquisition rate of the stimulus-response correspondence and a better RF estimation quality.

There are two major challenges in making advanced feature noise (or just *feature noise*) for reverse correlation: 1. A feature noise pattern must be presentable as a vector in Euclidean space. Or we cannot ensure the spherical symmetry of such a noise. 2. It can be converted back to basic visual feature in the same way as it can be detected from the basic features by the visual system of interest. Or we cannot ensure the advanced feature in the noise stimulus will be processed by the animal as expected. For example, the visual motion must be presented as a velocity vector whose phase angle and norm are corresponding to the moving direction and speed of the motion to ensure the spherical symmetry of the stimulus distribution[84]. Then the vector noise must be converted back to image pattern sequence by reversing the motion detection algorithm used by the visual system of the corresponding animal (e.g. the reversed Reichardt detector). Inevitably, the image patterns converted may contain other features besides visual motion, such as edge orientations and contrasts. And the motion detection is relied on the spatiotemporal structure of an image sequence, which means each vector noise pattern may require more than one image pattern to represent. These two issues should be aware in the construction of feature noise as the first issue may affect on the accuracy of RF estimation and the second directly affects the estimation efficiency.

Besides these challenges, it is also challenging to applying this method to calcium imaging data. In comparison with the neurophysiological activity recorded with electrophysiology method, the neural activity recorded with calcium imaging has three major disadvantages:

1. Poor SNR due to the relatively low amplitude of single-spike induced calcium activity and the photon shot noise in the calcium recording[8,83,96]

2. Slow decaying dynamics of calcium indicator[8,12] and low image acquisition rate

3. Low signal acquiring rate: although the sampling rate is largely depending on the imaging setups and configuration, in most cases it is not comparable with the sampling rate in electrophysiological recordings.

These problems pose challenges for the reverse-correlation based receptive field estimation which requires temporally precise and accurate detections of discrete neurophysiological events (e.g. spikes). Many of these challenges can be partially solved in the data analysis step by inferring spike activities from calcium signals[69,89]. However the precision of the spike inference is depending on the quality of calcium events. It is easier to detect and infer spikes for the strong calcium events induced by temporally clustered spikes than the weak events induced by sparse spikes. The feature noise should therefore have some degrees of temporal contiguity in order to trigger these temporally clustered spikes for robust receptive field estimation with noisy, slow calcium imaging data.

In conclusion, the reverse correlation technique provides an unbiased estimation of RF structure from the suprathreshold neural events under the *linearity* and the *spherical symmetry* assumption. This is critical for the compatibility with *in vivo* calcium imaging. And by constructing a spherical symmetric visual motion noise stimulus, the estimation efficiency problem in reverser correlation may be solved by using feature noise stimulus to a large extent, which should make it possible to transfer the success in the investigation of optic flow processing algorithm in Drosophila to vertebrate species.

3.5 LARVAL ZEBRAFISH AS A MODEL FOR OPTIC FLOW PROCESSING

To understand what algorithm is employed by vertebrates for optic flow processing, the animal model selected must show clear evidences for extracting the self-motion information from optic flow and utilizing this information in relevant behaviors (e.g. visual stabilization). And the identified or putative neural correlate of optic flow processing should be accessible for

systematic functional characterization with large-scale neurophysiological recordings at single cell resolution.

The optic flow processing in larval zebrafish has been studied since a few decades ago. The optic-flow related behaviors have been found and well characterized in zebrafish: The OKR and OMR are the innate behavior which contribute to gaze and body position stabilization in larval zebrafish respectively, and they can be robustly triggered with wide-field visual motion[41,49,62,64,67,70]. The OKR is the smooth rotational eye movement in the direction of visual motion perceived developed since 3 days post-fertilization (dpf)[16]. It is majorly triggered by rotational optic flow but also respond to translational optic flow[55]. The OMR in the previous studies are usually triggered by translational optic flow or wide-field visual motion with consistent moving direction, fish are swimming in the direction of optic flow or visual motion[67,75]. It can also be triggered with rotational optic flow[36] or even moving edge in the adjacent local region[41]. In most studies, the OMR and OKR behavior can be explained with an alternative hypothesis that they are not stabilizing the self-motion status but the retinal images. Therefore it is important to determine from the neural correlate level whether the neural correlate processed the self-motion induced optic flow represents self-motion information instead of retinal displacement fields.

The small and translucent brain of larval zebrafish and various established transgenic line allows systematic functional and anatomical characterization of the neural correlate for optic flow processing with calcium imaging or electron microscope (EM)[31,52,64]. The optic tectum and the pretectum are the two candidates for optic flow processing in larval zebrafish. The optic tectum receives direct input from retina and plays an important role in sensorimotor transformation of visual motion[57,67,75]. There is a large amount of direction selective neurons identified in the optic tectum which are mainly monocular tuned to four directions that are roughly correspond to the up/down/left/right direction in a 2D plane[34,57,92]. The ablation of optic tectum does not significant impair OKR or OMR, suggesting it does not play a major role in the sensorimotor transformation pathway for self-motion stabilization.

The pretectum, or area pretectalis (APT) also receives the directional information from retina via the arborization field (AF) 5[44]. The direction selective neurons in pretectum can be both monocular and binocular. Similar to the direction selective neurons in

optic tectum, the monocular direction selective pretectal neurons are also tuned to the four cardinal directions. The systematic characterization of the binocular optic flow sensitive neurons reveals a substantial amount of neurons in the pretectal anterior medial cluster (AMC) that are selective to binocular translational or rotational optic flow[49,92]. The characterization of the preferred direction selectivity of the binocular selective neurons in the left and right visual field suggests they might possess similar RF structure to certain self-motion induced optic flow fields[92]. In addition, another pretectal cluster has been found sufficient for inducing global optic flow-triggered OKR[98]. These evidences indicate the pretectum might play a critical role in the optic flow processing.

The clear optic flow-induced behavior, the identified neural correlate for optic flow processing, and the well-established functional and behavioral imaging tools make larval zebrafish a suitable animal model for the investigation of optic flow processing. Yet, the exact algorithm for self-motion induced optic flow processing remains unknown due to the lack of an efficient and accurate method for systematic, fine-scale motion RF characterization. Therefore my PhD study aims to:

1. Establish an efficient, robust method for fine-grained motion RF estimation in larvae zebrafish with calcium imaging.

2. Systematically characterize the RFs of motion sensitive neurons in the pretectal area.

3. Determine if these neurons may extract and encode the self-motion information from the self-motion induced optic flow with matched filter algorithm.

4. Predict the algorithmic and computational properties of optic flow processing based on the observations in the neural implementation, and verify these predictions with neurophysiological or behavioral experiments.

Part III

DISCUSSION

4

DISCUSSION

4.1 SPATIOTEMPORAL CORRELATION IN CMN

The CMN stimulus is not the first of its type, similar stimuli have been repeatedly invented and applied to reverse correlation since the early 90s[6,50,51,84]. The biggest difference between CMN and the other motion noise stimuli is the spatiotemporal contiguity which becomes its strength and weakness at the same time.

In the introduction I mentioned motion noise stimuli, as a feature noise stimuli must be converted back to image pattern sequences from velocity vectors to be perceived by animals. Therefore in the previous studies, each optic flow field pattern is usually presented with a short movie with fixed duration that is separated by static patterns during the inter-stimuli intervals. This phasic stimulus design limits this type of motion noise stimuli to only be used in the electrophysiology studies. Because only in these studies, the neural responses to transient motion during short stimulus phases can be well separated due to the fast kinetic of the spike signal[6]. For the neural signals with slow kinetic, the stimulus phases need to be extended to allow a reliable detection of neural responses to the optic flow pattern. This may lead to a longer recording time or fewer stimulus patterns to be displayed in a fix time period. For example, the motion RF estimation in a previous study took only 1.6 to 6.7 minutes to display 5000-20000 stimulus phase (20 milliseconds per stimulus phase)[6]. Due to the slow kinetic and the low sensitivity of the calcium signal, the stimulus phase should be at least 150 times longer (1 second stimulus displaying + 2 seconds interval) to trigger and separate calcium responses significantly higher than the baseline. The total recording time will thus be 4-16 hours which is not acceptable when considering the quality of neural responses during the long recording and the bleaching of the fluorophore in the calcium indicator.

The necessity of the spatiotemporal contiguity for efficient RF estimation

The temporal contiguity in the CMN stimulus allows the 1-to-1 conversion between the motion vector field and the binary movie frames (see the method section "Reichardt encoder" in the section 7.1 for details). Under the linearity assumption in the reverse correlation method, similar stimuli may evoke similar neural responses. The temporal contiguous in the CMN

may lead to the gradual change of the optic flow pattern. The temporal correlation in the stimulus should also be reflected in the neural responses. Therefore the neural responses to CMN should be temporally clustered which can be robustly detected in the calcium signal despite the low SNR and slow kinetic in comparison with electrophysiology signal.

The spatial contiguity is added for another reason. This RF estimation method was first developed for estimating the pretectal neurons selective to global instead of local visual motion. The previous study[49] and the paper in Appendix 7.2 show these neurons prefer spatially coherent wide-field motion. Based on this prior knowledge, we introduced the spatial contiguity to force the motion directions in local regions to be more spatially coherent to efficiently trigger their responses.

The strong correlation in the CMN stimulus may result in the underrepresentation of high frequency optic flow component

The spatiotemporal contiguity we introduced into the CMN stimulus is essentially strong spatiotemporal correlations which in principle should be avoided to preserve the spherical symmetry in the noise distribution that is essential for unbiased RF estimations[68]. However we may avoid misinterpreting the RF estimations if the biases are predictable or controllable. Unlike the correlation in the natural stimuli or pink noise distribution which affect the whole frequency spectrum, the correlation we introduced by convolving Gaussian white noise with a Gaussian kernel. The Gaussian noise is spherically symmetric. And the convolution operation may only dampen the power of high frequency components in the CMN[1]. And the distribution of the components with same frequency in the CMN may remain to be spherical symmetry. Therefore the only bias caused by the spatiotemporal correlation we introduced is the under-representation of the high frequency components in the RF estimation. In other words, the correlation reduced the sensitivity but not the accuracy of high frequency RF structure estimation. This is supported by the observations that both low and high frequency components such as the contraction/expansion point in some complex RFs estimated are similar to the corresponding structure in the self-motion induced optic flow they are tuned to (section 7.3).

1 The convolution operation is equivalent to the multiplication in the Fourier space. The Fourier transformation of the Gaussian function $\frac{1}{2\pi\sigma^2}\exp-\frac{x^2}{2\sigma^2}$ is another real-valued Gaussian function $\exp-\frac{k}{2}\sigma^2$. Therefore the multiplication will only alter the power but not the phase component of the Fourier transformed white noise

Due to this under-representation of high frequency temporal component, the RF estimation method using CMN is weak in estimating transient temporal RF structure which is hard to access anyhow due to the low calcium image acquisition rate (2Hz) and the slow calcium indicator kinetic (the decaying time is 1.7 second). Hence we chose not to further characterize the temporal structures of motion RF or to improve the spatial RF estimations by performing cross-correlation. In principle the temporal RF structure could be accessed by using the CMN with lower temporal contiguous radius and increasing the calcium imaging acquisition rate. The characterization of temporal RF structure may contribute to determine if and how the optic flow information is integrated in the temporal domain.

4.2 PROBLEMS IN THE REVERSE CORRELATION OF CALCIUM SIGNAL

The reverse correlation method, as mentioned before, estimates the stimulus-response function with the response-stimulus correspondence. An inaccurate response detection will result in poor quality and strong biases in the RF estimation. The responses of spiking neurons can be reliably detected as the voltage spikes with similar amplitude and certain refractory periods in electrophysiology recordings. The similar amplitude of these voltage spikes allows the spike detection by setting an adequate threshold, and the refractory period defined the minimum time bin for sampling and separating neural activities. In comparison with voltage spikes, the calcium indicators have much slower decaying kinetics which results in the convoluted fluorescent signal that are difficult to separate temporally. The nonlinear correlation between the fluorescent intensity of calcium indicator and the spike rate and the relatively poor SNR in most calcium recordings results in inaccurate inference of neural responses.

Sensitivity and precision problems in calcium event detection

To solve these problems, we binarize the calcium traces into a binary train of calcium events, which are the strong neural responses that lead to abrupt and significant increases of calcium activities. And these calcium events are detected by a gaussian-mixture model (GMM) based algorithm. This method is compatible with the low temporal sampling rate and SNR in our calcium recording. However it has two obvious problems: 1) it only detects the strong and/or temporally clustered neural responses occurred before the over-saturation of the fluorescent signal, and 2) the binarization erase the amplitude information of calcium events which reflects the difference in response inten-

sity. The loss of stimulus-response correspondence information resulted by these problems may lead to a lower estimation quality and under-representation of the RF components with less weight. Another problem is the event detection precision. The false positive rate of our event detection algorithm tested with *in silico* data is relatively low (5%). However, if we calculate the detection precision as the percentage of the true positive events in the total number of events detected, the in the neurons generates few or no calcium events (event rate <5%), the detection precision will be lower than 50%.

This problem is partially solved by the two-step nonparametric cluster-based bootstrapping (NCB) test we developed as it can well control the false positive rate in the final RF estimation to be under 5%[59,66]. However the two-step NCB test also has the precision problem: if none of the neurons in the neural correlate are sensitive to motion, the statistic test will still find 5% neurons with motion RFs. These problems are all due to the lack of prior knowledge of the existence of stimulus-response correspondence. A potential solution is to display a strong motion stimulus before displaying the CMN stimulus, so we may determine if a neuron is motion sensitive based on its responses to this stimulus. But the paper in Appendix 7.2 suggests it is difficult to find an "omnipotent" stimulus to trigger all motion sensitive neurons' responses as they may tune to the visual motion features that are mutually exclusive. For example, some neurons respond only to local visual motion which are insufficient to trigger strong responses in the neurons prefer global motions. Thus we might exclude some potential motion sensitive neurons we are not expecting with this method. Another less biased way is to record the responses to multiple different CMN stimulus and evaluate the reproducibility of the RF estimations. The real RF structure should be invariant in all RF estimations of the same neuron if the CMN stimuli used are all sufficiently long. This method is equivalent to estimate RFs with a longer CMN stimulus. The long recording time may lead to other problems in performing calcium experiments, such as large motion artifact and strong bleaching effect in the calcium recordings, which should be balanced with the RF estimation quality.

The statistic power of the two-step NCB test is limited to the cluster level but not single unit level, which means the significant spatial cluster it identified from the raw RF estimation may also contain noise[59,66]. And conversely, there might also be significant RF units in the insignificant cluster, especially the

scattered ones which are not expected by the statistic test which assumes the spatial contiguity of RF structures. The significance threshold in the first step of the NCB test governs both the single unit and cluster-level type I error. A low significance threshold will make the statistic test be more sensitive to the RF structure assigned with less weight, but it will also make it harder to locate the real RF structure in the significant cluster identified in the second step. While increasing this threshold may make it easier to locate the real RF structure in the significant cluster, but it may also result in a higher false negative rate at both the single unit and cluster level. The sensitivity, localization power and the false positive rate of this two-step NCB test form a vicious triangle. In our study, the false negative error is as important as the false positive error. For example, the omission may result in the misidentification of unimodal RFs which are actually bimodal RFs. Thus in the paper in Appendix 7.3, we traded the localization power with the sensitivity, which make the method weak in determine the precise RF size and shape.

The information acquisition rate is the limiting factor for RF estimation precision, sensitivity and accuracy

In the end, I would like to discuss the limitation of reverse correlation itself. As mentioned in the introduction, there are two major assumptions for the reverse correlation technique: the *spherical symmetry* assumption of the noise stimulus, and the *linearity* assumption of the RF structure. If the second assumption may not be satisfied, the reverse correlation[2] may not recover the real RF structures at all in the extreme case. For example, the RF subunits of the complex cells in Macaque V1 may disrupt the RF estimation by reverse correlation[78,91]. Another common case is the identification of the suppression RF component in the neurons received static inhibitory modulation[79]. In these cases, the subunit information may to some extent be determined with spike-triggered covariance analysis[68,78–80]. It is less likely that the complex RFs for self-motion induced optic flow detection may have multiple overlapping excitatory subunits (i.e. the RF components with same spatial location but different direction preferences) if they implement the *identity* principle of self-motion induced optic flow processing. However it is possible that there may be some suppressive RF components to enhance the precision of self-motion status estimation. Another reason these suppressive components may be problematic for reverse correlation estimation is their effects may be mixed with the speed tuning properties (see Figure 7 in Appendix 7.1 and Figure 4 in Appendix 7.3). For a neuron tuned to the motion

The limitation of reverse correlation in estimating RFs with subunits or nonlinear RFs

2 The reverse correlation method here includes only the spike-/event-triggered average

speed $\dot{s}$ that $R(\dot{s}) > R(x)$ for $x \in \mathbb{R}^+$, its speed tuning function $R(s)$ has:

$$R(a\dot{s}) <= R(\dot{s}) < aR(\dot{s}) \qquad \textit{(violation of homogeneity)}$$
$$R(\dot{s}+b) <= R(\dot{s}) < R(\dot{s}) + R(b) \qquad \textit{(violation of additivity)}$$

where a, b and $R(b) > 0$. The homogeneity and the additivity are basic properties of any linear system. The violation of these two properties suggests the speed-related terms in the responses function may not be estimated with reverse correlation method.

If a motion RF has suppressive components, and it also tuned to speed, the nonlinearity in speed tuning may ravel with the suppressive components to make the reverse correlation estimation into a nightmare. Normally a suppressive component may lead to a reduction in the response intensity to the increase of stimulus intensity. However the nonlinearity in the speed tuning will make this relationship hard to measure because of the lack of good measure of stimulus intensity. In Appendix 7.1, we proposed a low speed range to circumvent this problem. All visual motion sensitive neurons by definition should respond to visual motion with non-zero speed. If we assume the speed tuning function of all motion sensitive neurons are smooth and continuous, then there should be a speed range $(0, q]$, in which the speed function is monotonically increased, for unraveling the excitatory-suppressive index and the speed tuning properties. But it still remains to be a problem to determine this speed range without making any assumption.

To solve this problem, it is necessary to review the spherical symmetry assumption. It does not pose any constrain in the distribution of the norm of stimulus vectors which are associated with the visual motion speeds. The Gaussian white noise distribution used in our studies for example, may over-represent of the low-speed visual motion and under-represent the high-speed ones. This can be fixed by replacing the Gaussian white noise with the noise stimulus sampled from a uniform spherical distribution (e.g. Figure 3.5b). The nonlinear speed tuning property of the neuron may then be estimated as RF weights with reverse correlation[3].

The speed tuning problem may get more complicated in the translation sensitive neurons we found in the paper in Appendix

3 Again this method only works for the RFs with linear orthogonal presentations in the stimulus parameter space. It may fail to estimate those nonlinear RFs, for instance, a speed-tuning RF with no direction selectivity.

7.3. They are more likely tuned to the speed of self translation instead of the local motion speed. But the former can only be inferred with the spatial integration of the latter which are conjugated with the depth information of the visual cue as indicated in the equation 3.4. Also the animal must determine if the motion represented by each visual cue is induced by self-motion or object motion before the integration for speed tuning. Since these questions are not related only with RF estimation, I will discuss them more in the next section about matched filter algorithm.

4.3 ENCODING QUALITY OF MATCHED FILTER ALGORITHM

In the introduction I discussed the advantages and disadvantages of the representative algorithms proposed for optic flow processing. These discussions would be pointless if none of them are actually employed by any animal to process self-motion induced optic flow. The third paper in Appendix 7.3 shows that the pretectal neurons with complex RFs served as linear filters for self-motion induced optic flow processing in larval zebrafish. These evidences indicate the matched filter algorithm is implemented in zebrafish larvae for self-motion extraction and encoding. Now it may be worth to discuss which factors in this neural implementation may affect the algorithm performance in terms of the encoding quality of self-motion information.

The evaluation of encoding quality of self motion information involves at least three factors: code ambiguity, noise robustness and decodability. These factors are often tangled up at the implementational level. The code ambiguity is related with the decomposability and the identity of translation and rotation information. As shown in the equations 3.9 and 3.11, the decomposability is implemented in the matched filter model by the distributive law of the inner product. However this requires the visual cue to be spherically symmetric distributed at the same depth, which is not common in natural scenes. It also requires the receptive field to sample the entire visual field, while the complex RFs of the pretectal neurons only cover some parts of the visual field. Thus the decomposability in this neural implementation of the matched filter algorithm need to be re-evaluated. For a complex RF $\vec{V}$ sampling n locations, its re-

sponse r to the optic flow with a translational component $\vec{F_T}$ and a rotational component $\vec{F_R}$ is:

$$\begin{aligned} r &= \sum \vec{V} \cdot (\vec{F_T} + \vec{F_R}) \\ &= \sum_{i=1}^{n} \underbrace{\vec{v_i} \cdot \vec{f_{T_i}}}_{\text{trans. scr.}} + \underbrace{\vec{v_i} \cdot \vec{f_{R_i}}}_{\text{rot. scr.}} \end{aligned}$$

The variation of decomposability across the visual field is resulted by the geometry of rotation- and translation-induced optic flow fields

There are at least 2 points in any combinations of translation- and rotation-induced optic flow field where $\vec{f_T}$ is parallel to $\vec{f_R}$ [4]. The translation score (trans. scr.) and rotation score (rot. scr.) at these locations will be similar and thus contribute very little to the overall code decomposability regardless what value $\vec{v_i}$ has. On the plane determined by the self-translation and self-rotation vectors, all $\vec{f_T}$ are perpendicular to $\vec{f_R}$, including points on this plane in the RFs may increase the decomposability. For a set of self-motion induced optic flow field whose self-rotation axes are distributed in a spherically uniform way, the probability density of these zero-decomposability points are high near the FoC and focus of expansion (FoE) in the optic flow fields but lower near the plane passing the origin (the observer) and perpendicular to the corresponding self-translation direction. This is consistent with the observation in our study (Figure 4E in the section 7.3) that the mode centers of the bimodal complex RFs are distributed near the planes perpendicular to the translation direction. This implies these neurons may maximize the decomposability of the representation of self-translation and self-rotation by selectively integration of visual motion information from different parts of the visual field. Again, to confirm this, we must exclude the potential biases in the RF estimation protocol. And since in this study the visual stimulation setup does not cover the entire visual field, we didn't have the complete RFs estimated which may be essential to determine if they implement the optimal sampling strategy for optic flow processing with matched filter model.

The translation-rotation decomposability is affected by the parts of optic flow field RF covered

The identity principle is partially implemented in the structure of matched filter, as the structure of each matched filter is corresponding to a certain type of self-motion induced optic flow fields[5]. Since the matched filter implemented in zebrafish

4 The analytical solution is too long to include. But it consists of two terms which are the sum and the cross product of the unitary self translation and rotation velocity vectors. The weight of these two terms can be described by a function of the angle between the translation and rotation vectors

5 As the equations 3.9 and 3.11 showed that it requires at least 3 translational and 3 rotational matched filter whose preferred translation/rotation axes are not coplanar, to determine the precise translational and rotational self-motion vectors. Therefore the identity principle is implemented by both the linear filter structure and the decoding of the matched filter codes

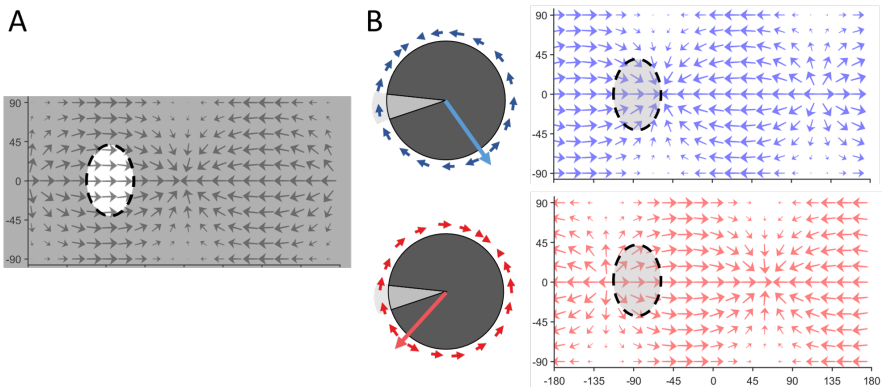


Figure 4.1: **A** An unimodal RF with the same structure as the highlighted areas in the forward-translation-induced optic flow. **B** The two optic flow fields induced by self-translation in two different directions will trigger identical responses in the neuron with the unimodal RF in A because of the inner products of each vector pair in the sampled region are identical.

does not sample the entire visual field, the spatial locations sampled by the complex RFs may also affect the implementation of the identity principle, which in turns affect the code ambiguity. An example is shown in the Figure below:

The implementation of the identity principle can be improved by increasing the size of the RF as shown in a previous study[10]. It may also be improved by splitting the RF into two components to increase the RF span without changing the total RF areas as shown in the figure 4.2[6].

The noise robustness of the matched filter algorithm has been discussed in the paper in Appendix 7.3, but mostly at the population instead of single neuron level. The noise robustness at the single RF level is often tangled with the code ambiguity as they are all related with the implementation of the identity principle. There are two types of noise in the self-motion induced optic flow fields: one is resulted from the inaccurate detection of visual motion, the other is the independent object motion that contaminates the self-motion induced optic flow field. The former mostly affects the algorithms requiring high precision in the visual motion detection, such as the analytical solution proposed by Longuet-Higgins and Prazdny[56]. The latter may affect all algorithms determining the self-motion information

The first noise is in the implementation of visual motion detection and thus affects all visual motion processing, the second is specific to optic flow processing

6 Note if I increase the span of the bimodal RF in Figure 3.2 to 180 degree, it may again respond to the two optic flow fields in the same way because of the symmetry in the translation-induced optic flow fields. This may explain that we found fewer bimodal neurons with intermodal distances equal to 180 in the third paper in Appendix 7.3

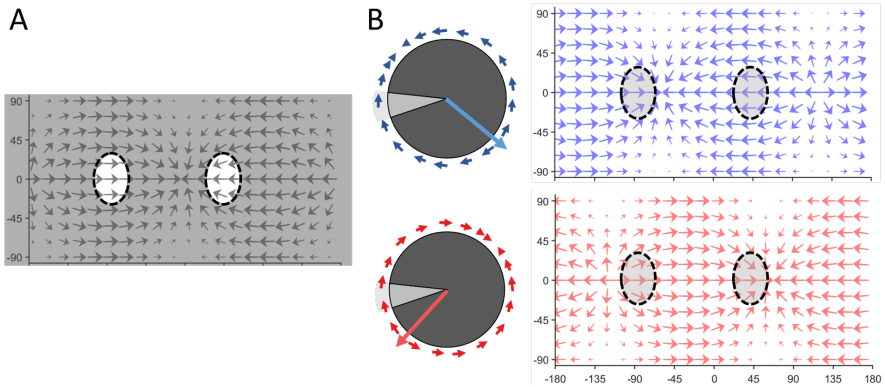


Figure 4.2: **A** An bimodal RF with the same structure as the highlighted areas in the forward-translation-induced optic flow. **B** The two translation-induced optic flow field now may trigger different responses due to the structure difference in the region covered by the RF components in the right visual field.

solely from instantaneous optic flow. In naturalistic scenes, the distribution of both types of noise are spatially correlated, especially the second type of noise[5]. This means if a part of an optic flow field is corrupted by the second type of noise, the neighboring area are more likely to also be affected. Therefore for the RFs with the same size, the more scattered RFs may be more robust to the noise induced by object motion. However this also means the size of each spatially connected component will be smaller which may reduce the visual motion detection accuracy and robustness to the first type of noise. It is possible that the balance between the code robustness to these two types of noise resulted most complex RFs we found are bimodal but not unimodal or trimodal.

Besides the code ambiguity and the noise robustness, the decodability is also a major factor for judging the encoding quality of self-motion information. The decodability describes the computational complexity of the decoding process for accessing the self-motion information from the population code. For example if the self-translation direction is encoded by n neurons with the "one-shot" algorithm in the Dahmen et al. paper[10] (equivalent to the one-hot code), the complexity will be $O(n)$. And to decode the self-motion information encoded by the matched filters, the computational complexity will be $O(n^k)$, $k > 2$ as the decoding algorithm must involve matrix multiplication[7][10,18].

7 The complexities calculated here are not very precise as the computation happened in individual neuron or neural circuit level are highly parallelized.

I consider the decodability as a major factor for encoding quality of self-motion information because the computational complexity may be also reflected in the neural implementation. Increasing amount of information will be loss during the signal computation and transmission in the dendritic and neural circuits implementing a highly complicate decoding algorithm[8]. In the optic flow processing case, the self-motion information can be decoded from the population code of pretectal neurons with complex RFs, or the population code of the direction selective neurons with unimodal RFs, however the latter is more ambiguous in terms of the translation-rotation decomposability which requires extra decoding step to solve. The decodability of the unimodal RFs are thus lower than the one of complex RFs.

4.4 THE DOWNSTREAM CIRCUIT IN THE SENSORIMOTOR TRANSFORMATION OF OPTIC FLOW

The decomposability and the identity principles can only apply to the encoding part of self-motion induced optic flow processing. Because once the self-motion is retrieved from the optic flow, the computational goal for the downstream circuits in this sensorimotor transformation will be changed. One possible goal of the downstream circuit is to transform the self-motion information to motor command for generating behavioral responses to optic flow. In this case, the code (un)ambiguity, noise robustness and decodability should be evaluated with a completely different standard. For example, If the fish in the figure below can only swim in the four directions indicated with the black arrows, the decodability of the two self-motion representations in blue will be higher than the other representation in red, although they are mathematically equivalent to each other, because the sensorimotor transformation of self-motion representation does not involve matrix multiplication which is required by the other representation.

Figure 4.3 is a simplified illustration for the "mode sensing" hypothesis which hypothesized the basis for representing self-

They are just used for demonstrating the difference of the decodability of different encoding algorithms

8 This might also contribute to answer why feature extraction is necessary: If we ignore this factor, then there is nothing wrong to implement each sensorimotor transformation with an independent neural circuit which may not require to extract any feature. However the overall computational complexity of such systems will be much higher than the systems extract common features in the sensory processing which can be shared by many downstream process.

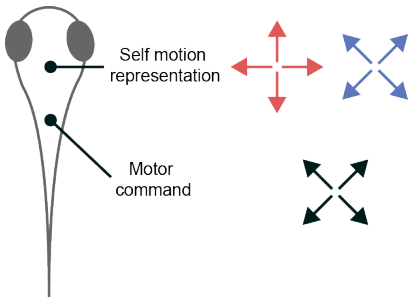


Figure 4.3: The red and blue arrows showed two sets of basis function which are mathematically identical in representing the self-translation information extracted from optic flow. The black arrows indicate the consequential self-translation direction for each elemental motor commands. The sensory representation of self-translation matched the best with the representation of motor commands (blue arrows) produced the most efficient sensorimotor transformation in this simplified circuit.

The sensory representation should also reflect the computational principle and algorithm employed by the downstream pathway of sensorimotor transformation in the motor systems

motion status should be the same as the basis for representing all possible motor states (or "motor mode") to maximize the sensorimotor transform efficiency or the decodability of sensory representation[48,88]. Because then the population code of self-motion can be directly used as motor commands without any further transformation[48]. This is supported by the non-uniform distribution of the preferred rotation axes of LPTC neurons which servers as matched filter for optic flow detections in flies[88] and also the distribution of preferred translation direction of the translation sensitive pretectal neurons in our study (section 7.3). However this hypothesized algorithm assumes the optic-flow related sensorimotor transformation is mostly linear, which might be oversimplified and remained to be confirmed.

The controversies about the roles of pretectum and hindbrain in the sensorimotor transformation of global optic flow

In larval zebrafish, the research of sensorimotor transformation of optic flow are mainly associated with two anatomical structures: the pretectum and the hindbrain. Their roles in the optic flow related sensorimotor transformation are somewhat controversial in previous studies. Some suggested the transformation of optic flow to self-motion information is conducted only by the pretectum, which is consistent with the finding in the paper in Appendix 7.2[49,92]. While some other study claimed not the pretectum, but also the hindbrain are involved in the extraction of self-motion information[64]. As I have already discussed in the paper in Appendix 7.3, the controversial explanations may be explained by the distinct optic flow statistic in the two different regions of the visual field stimulated in these studies. By taking this difference into account, these controversial observations can be well explained with the matched filter model.

Another pair of seemingly controversial findings argues the role of pretectum in generating motor responses to optic flow: one study reported strong correlation between pretectum and the visual stimulation induced turning behavior[102], while the other study suggests the hindbrain but not the pretectum is responsible for making decision of turning behavior[1]. As suggested in the second study, it is possible the hindbrain is involved in the temporal integration of self-motion information extracted by pretectum from instantaneous optic flow to prevent making frequent motor responses which are expensive in terms of energy expenses.

4.5 GAPS BETWEEN RFS, MATCHED FILTERS, AND BEHAVIOR

As I stated at the beginning of this thesis that I consider my PhD project as a showcase of Marr's three levels of understanding framework. In the introduction I have mentioned the two computational principles of self-motion induced optic flow processing. I discussed how the matched filter algorithm may recover the self-motion status by satisfying these two principles, and why the RF estimation may reflect the neural implementation of matched filter algorithm. This is confirmed by the results in the section 7.3, and we found both physiological and behavioral evidence that confirms one of the computational principles, the decomposability is implemented in the optic flow processing system of larval zebrafish. In the discussion I have discussed how the neural implementation of matched filter algorithm reflected by their RF structures may affect the algorithm performance in terms of the code ambiguity, noise robustness and decodability. I very much wish these results and discussion can make a complete demonstration of the three-level framework, however I must admit there are at least two major gaps in this study. In this last section I will discuss how could these gaps be filled in the future studies.

This gap is mainly due to the lack of systematic characterization at the implementational levels

The first gap is in between the algorithmic and implementational level. We do not know how these complex RFs are implemented in the neural correlates. Are they implemented by the dendritic field of the corresponding neurons or inherited from the upstream neurons. We can neither answer what role they may play in the downstream circuits, for example, how may the self-motion information be decoded and use the self-motion information in behavior. Systematic characterization of the structural and functional connectome of the neural correlates

involved in the optic flow processing may be desired for filling this gap. By revealing the structural connectome of AMC neurons, we may identify the possible upstream and downstream neural correlates. The characterization of the upstream pathway and the dendritic anatomy may contribute to explain the emergence of the matched filter algorithm. Characterization of downstream neural correlate may update our understanding of the role of AMC in the whole sensorimotor transformation pathway. It might also contribute to the understanding of the algorithm: why it is necessary and how do the properties of the matched filter algorithm limit the choices of possible algorithms and the computational goal of the downstream neural correlates. Currently there are two ways to obtain the full structural connectome: the high-resolution serial-section EM[31] and the whole-brain scale imaging of neurons sparsely expressing GFP[52]. The former technique allows to acquire the complete structural connectome information in the whole brain. However to my knowledge there is no fully automated methods at the current stage for efficiently retrieving the connectome information from EM datasets. Another challenge is to identify functional cluster from the EM data, which require accurate alignment of the brain structure in the EM dataset with the brain structures obtained from fluorescence microscope imaging. These problems to some extent can be solved with the sparse labeling approach which is compatible with calcium imaging and also relatively easy for automatic tracing[52]. However the sparse labelling also results in the structural information acquired from one fish are not complete, and it usually requires to merge the labelled data from multiple fish to draw the conclusion. In such cases, the individual variance must be taken into account for data interpretation.

The structural connectivity may only provide a rough outline of the information flow framework, as it tells very little about the dynamic functional components. For example, the conditional connectivity as discussed in detail in a recent review[19]. Thus the functional connectivity must also be evaluated for filling the gap between the understanding at the implementational and algorithmic level. Most methods for evaluating large-scale functional connectivity are based on the correlation of neural responses and the stimulus or motor output[13,58,102]. The observation made in these studies are valuable in forming hypotheses about the possible algorithms and computational goal, or predict the functional and structural connectivity which can be verified in the optogenetic or EM studies. However these observations should not be considered as solid evidences for

revealing functional connectivity especially in those large-scale neurophysiological recordings where many "significant" neural correlates may emerge due to the lack of multiple testing correction.

The other gap in my study is somewhat related to my criticism of the functional correlation as a method for revealing functional connectivity above. In our study we implied that the decomposition of translational and rotational optic flow component revealed in the behavioral experiments is accomplished by the matched filter algorithm implemented in the AMC neurons with complex RFs. This is weakly supported by the invariant responses of the translation sensitive neurons in the AMC to the translation-induced optic flow field with or without rotational interferences. Although this evidence is insufficient to prove these neurons are the neural correlate for the optic flow decomposition, the identification of the topographical map of self-translation direction in the AMC and the behavioral responses to optic flow mixture in our study allows future studies to further investigate their relationship in a quantitative way. For instance, one may manipulate the activities of the AMC neurons preferring a certain translation directions with laser ablation, to test if and how may this affect the decomposition of certain combinations of translational and rotational optic flow in OKR and OMR. It is also possible that zebrafish may have more than one neural correlates implementing the matched filter algorithm. Or there may be more than one algorithm employed by larval zebrafish for optic flow processing. It would be curious to know if there is any alternative algorithm for optic flow processing when the matched filter algorithm may not determine self-motion status from some specific optic flow such as the "optic flow illusion". This stimulus is constructed by overlaying two sets of motion cues which are corresponding to translational and rotational optic flow respectively. In human, the perceived translation direction is shifted by the rotational optic flow components[77].

In the end, I would like to very briefly talk about an issue in broadcasting Marr's framework to the subcomponents of the visual systems (e.g. neural correlates or individual neurons). The definition of the three levels may become ambiguous when applying to the subcomponents of the system. For example, the computational goal of each individual translation-sensitive neuron may not be the same as the goal of the entire system which is to extract self-motion information as it is only available in the population code instead of the responses of individual

neurons. And the computational goals may also vary across the stages of visual processing and neural correlates. However in these cases, the core of the three-level framework is still valid. The understanding of the algorithms can be improved by investigating the computational goal and principles or revealing the neural implementation. The investigation of computational goal may clarify the visual feature the components of interest responding to, and thus narrow down the possible functional role of these components. And the investigation of the neural implementational may narrow down their functional role and suitable putative algorithms. The investigation in these two directions may help to answer where do the functions of this component emerge from, and why it has such a function. These answers may eventually be summarized into an algorithm, that allows us to understand and most importantly, to predict the behavior of the system or the neural ensemble of interest.

Part IV

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Bahl, A. & Engert, F. Neural Circuits for Evidence Accumulation and Decision Making in Larval Zebrafish. *Nature Neuroscience* **23,** 94–102. ISSN: 1546-1726 (Jan. 2020).
2. Barlow, H. B. in *Sensory Communication* (ed Rosenblith, W. A.) 216–234 (The MIT Press, Sept. 2012). ISBN: 978-0-262-51842-0.
3. Bhagavatula, P. S., Claudianos, C., Ibbotson, M. R. & Srinivasan, M. V. Optic Flow Cues Guide Flight in Birds. *Current Biology* **21,** 1794–1799. ISSN: 0960-9822 (Nov. 2011).
4. Bishop, L. G., Keehn, D. G. & McCann, G. D. Motion Detection by Interneurons of Optic Lobes and Brain of the Flies Calliphora Phaenicia and Musca Domestica. *Journal of Neurophysiology* **31,** 509–525. ISSN: 0022-3077, 1522-1598 (July 1968).
5. Black, M. J. Robust Incremental Optical Flow, 280.
6. Borghuis, B. G., Perge, J. A., Vajda, I., van Wezel, R. J. A., van de Grind, W. A. & Lankheet, M. J. M. The Motion Reverse Correlation (MRC) Method: A Linear Systems Approach in the Motion Domain. *Journal of Neuroscience Methods* **123,** 153–166. ISSN: 0165-0270. pmid: 12606064 (Mar. 2003).
7. Brockerhoff, S. E., Hurley, J. B., Janssen-Bienhold, U., Neuhauss, S. C., Driever, W. & Dowling, J. E. A Behavioral Screen for Isolating Zebrafish Mutants with Visual System Defects. *Proceedings of the National Academy of Sciences* **92,** 10545–10549. ISSN: 0027-8424, 1091-6490 (Nov. 1995).
8. Chen, T.-W. *et al.* Ultrasensitive Fluorescent Proteins for Imaging Neuronal Activity. *Nature* **499,** 295–300. ISSN: 1476-4687 (July 2013).
9. Chichilnisky, E. J. A Simple White Noise Analysis of Neuronal Light Responses. *Network: Computation in Neural Systems* **12,** 199–213. ISSN: 0954-898X. pmid: 11405422 (Jan. 2001).
10. Dahmen, H.-J., Franz, M. O. & Krapp, H. G. in *Motion Vision: Computational, Neural, and Ecological Constraints* (eds Zanker, J. M. & Zeil, J.) 143–168 (Springer, Berlin, Heidelberg, 2001). ISBN: 978-3-642-56550-2.

11. Dahmen, H.-J., Wüst, R. M. & Zeil, J. Extracting Egomotion Parameters from Optic Flow: Principal Limits for Animals and Machines. *From living eyes to seeing machines* (1997).

12. Dana, H. *et al.* High-Performance Calcium Sensors for Imaging Activity in Neuronal Populations and Microcompartments. *Nature Methods* **16,** 649–657. ISSN: 1548-7105 (July 2019).

13. Diaz Verdugo, C. *et al.* Glia-Neuron Interactions Underlie State Transitions to Generalized Seizures. *Nature Communications* **10,** 3830. ISSN: 2041-1723 (Aug. 2019).

14. Douglas, R. M., Neve, A., Quittenbaum, J. P., Alam, N. M. & Prusky, G. T. Perception of Visual Motion Coherence by Rats and Mice. *Vision Research* **46,** 2842–2847. ISSN: 0042-6989 (Sept. 2006).

15. Drews, M. S., Leonhardt, A., Pirogova, N., Richter, F. G., Schuetzenberger, A., Braun, L., Serbe, E. & Borst, A. Dynamic Signal Compression for Robust Motion Vision in Flies. *Current Biology* **30,** 209–221.e8. ISSN: 0960-9822 (Jan. 2020).

16. Easter, S. S. & Nicola, G. N. The Development of Eye Movements in the Zebrafish (Danio Rerio). *Developmental Psychobiology* **31,** 267–276. ISSN: 0012-1630 (Dec. 1997).

17. Eichner, H., Joesch, M., Schnell, B., Reiff, D. F. & Borst, A. Internal Structure of the Fly Elementary Motion Detector. *Neuron* **70,** 1155–1164. ISSN: 08966273 (June 2011).

18. Franz, M. O. & Krapp, H. G. Wide-Field, Motion-Sensitive Neurons and Matched Filters for Optic Flow Fields. *Biological Cybernetics* **83,** 185–197. ISSN: 1432-0770 (Aug. 2000).

19. Friedrich, R. W. & Wanner, A. A. Dense Circuit Reconstruction to Understand Neuronal Computation: Focus on Zebrafish. *Annual Review of Neuroscience* **44,** 275–293 (2021).

20. Geiger, G. & Nässel, D. R. Visual Orientation Behaviour of Flies after Selective Laser Beam Ablation of Interneurones. *Nature* **293,** 398–399. ISSN: 1476-4687 (Oct. 1981).

21. Gibson, J. J. The Visual Perception of Objective Motion and Subjective Movement. *Psycho/. Rev,* 304–314.

22. Gibson, J. J. *The Perception of the Visual World* xii, 242 (Houghton Mifflin, Oxford, England, 1950).

23. Goller, B. & Altshuler, D. L. Hummingbirds Control Hovering Flight by Stabilizing Visual Motion. *Proceedings of the National Academy of Sciences* **111,** 18375–18380. ISSN: 0027-8424, 1091-6490 (Dec. 2014).

24. Gronenberg, W., Milde, J. J. & Strausfeld, N. J. Oculomotor Control in Calliphorid Flies: Organization of Descending Neurons to Neck Motor Neurons Responding to Visual Stimuli. *Journal of Comparative Neurology* **361,** 267–284. ISSN: 1096-9861 (1995).

25. Haag, J., Denk, W. & Borst, A. Fly Motion Vision Is Based on Reichardt Detectors Regardless of the Signal-to-Noise Ratio. *Proceedings of the National Academy of Sciences* **101,** 16333–16338. ISSN: 0027-8424, 1091-6490 (Nov. 2004).

26. Haag, J., Wertz, A. & Borst, A. Integration of Lobula Plate Output Signals by DNOVS1, an Identified Premotor Descending Neuron. *Journal of Neuroscience* **27,** 1992–2000. ISSN: 0270-6474, 1529-2401 (Feb. 2007).

27. Hassenstein, B. & Reichardt, W. Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers Chlorophanus. *Zeitschrift für Naturforschung B* **11,** 513–524. ISSN: 1865-7117, 0932-0776 (Oct. 1956).

28. Hausen, K., Wehrhahn, C. & Boycott, B. B. 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,** 211–216 (Sept. 1983).

29. Hausen, K. & Egelhaaf, M. Neural Mechanisms of Visual Course Control in Insects. *Facets of vision* (1989).

30. Hengstenberg, R. Gaze Control in the Blowfly Calliphora: A Multisensory, Two-Stage Integration Process. *Seminars in Neuroscience. Sensori-Motor Integration* **3,** 19–29. ISSN: 1044-5765 (Feb. 1991).

31. Hildebrand, D. G. C. *et al.* Whole-Brain Serial-Section Electron Microscopy in Larval Zebrafish. *Nature* **545,** 345–349. ISSN: 1476-4687 (May 2017).

32. Horseman, B. G., Macauley, M. W. S. & Barnes, W. J. P. Neuronal Processing of Translational Optic Flow in the Visual System of the Shore Crab Carcinus Maenas. *Journal of Experimental Biology* **214,** 1586–1598. ISSN: 0022-0949, 1477-9145. pmid: 21490266 (May 2011).

33. Hubel, D. H. & Wiesel, T. N. Receptive Fields of Single Neurones in the Cat's Striate Cortex. *The Journal of Physiology* **148,** 574–591. ISSN: 1469-7793 (Oct. 1959).

34. Hunter, P. R., Lowe, A. S., Thompson, I. D. & Meyer, M. P. Emergent Properties of the Optic Tectum Revealed by Population Analysis of Direction and Orientation Selectivity. *Journal of Neuroscience* **33,** 13940–13945. ISSN: 0270-6474, 1529-2401 (Aug. 2013).

35. Huston, S. J. & Krapp, H. G. Visuomotor Transformation in the Fly Gaze Stabilization System. *PLOS Biology* **6,** e173. ISSN: 1545-7885 (July 2008).

36. Imada, H., Hoki, M., Suehiro, Y., Okuyama, T., Kurabayashi, D., Shimada, A., Naruse, K., Takeda, H., Kubo, T. & Takeuchi, H. Coordinated and Cohesive Movement of Two Small Conspecific Fish Induced by Eliciting a Simultaneous Optomotor Response. *PLOS ONE* **5,** e11248. ISSN: 1932-6203 (June 2010).

37. Joesch, M., Schnell, B., Raghu, S. V., Reiff, D. F. & Borst, A. ON and OFF Pathways in Drosophila Motion Vision. *Nature* **468,** 300–304. ISSN: 1476-4687 (Nov. 2010).

38. Jonas, E. & Kording, K. *Could a Neuroscientist Understand a Microprocessor?* tech. rep. biorxiv;055624v1 (May 2016).

39. Jr, W. H. W., Kay, B. A., Zosh, W. D., Duchon, A. P. & Sahuc, S. Optic Flow Is Used to Control Human Walking. *Nature Neuroscience* **4,** 213–216. ISSN: 1546-1726. pmid: 11175884 (Feb. 2001).

40. Karmeier, K., Krapp, H. G. & Egelhaaf, M. Robustness of the Tuning of Fly Visual Interneurons to Rotatory Optic Flow. *Journal of Neurophysiology* **90,** 1626–1634. ISSN: 0022-3077 (Sept. 2003).

41. Kist, A. M. & Portugues, R. Optomotor Swimming in Larval Zebrafish Is Driven by Global Whole-Field Visual Motion and Local Light-Dark Transitions. *Cell Reports* **29,** 659–670.e3. ISSN: 2211-1247 (Oct. 2019).

42. Koenderink, J. J. & van Doorn, A. J. Facts on Optic Flow. *Biological Cybernetics* **56,** 247–254. ISSN: 1432-0770 (June 1987).

43. Koenderink, J. J. Optic Flow. *Vision Research* **26,** 161–179. ISSN: 0042-6989 (Jan. 1986).

44. Kramer, A., Wu, Y., Baier, H. & Kubo, F. Neuronal Architecture of a Visual Center That Processes Optic Flow. *Neuron* **103,** 118–132.e7. ISSN: 0896-6273 (July 2019).

45. Krapp, H. G., Hengstenberg, B. & Hengstenberg, R. Dendritic Structure and Receptive-Field Organization of Optic Flow Processing Interneurons in the Fly. *Journal of Neurophysiology* **79,** 1902–1917. ISSN: 0022-3077. pmid: 9535957 (Apr. 1998).

46. Krapp, H. G. & Hengstenberg, R. Estimation of Self-Motion by Optic Flow Processing in Single Visual Interneurons. *Nature* **384,** 463. ISSN: 1476-4687 (Dec. 1996).

47. Krapp, H. G. & Hengstenberg, R. A Fast Stimulus Procedure to Determine Local Receptive Field Properties of Motion-Sensitive Visual Interneurons. *Vision Research* **37,** 225–234. ISSN: 0042-6989 (Jan. 1997).

48. Krapp, H. G., Taylor, G. K. & Humbert, J. S. in *Frontiers in Sensing: From Biology to Engineering* (eds Barth, F. G., Humphrey, J. A. C. & Srinivasan, M. V.) 101–114 (Springer, Vienna, 2012). ISBN: 978-3-211-99749-9.

49. Kubo, F., Hablitzel, B., Dal Maschio, M., Driever, W., Baier, H. & Arrenberg, A. B. Functional Architecture of an Optic Flow-Responsive Area That Drives Horizontal Eye Movements in Zebrafish. *Neuron* **81,** 1344–1359. ISSN: 0896-6273 (Mar. 2014).

50. Kühn, N. K. & Gollisch, T. Joint Encoding of Object Motion and Motion Direction in the Salamander Retina. *Journal of Neuroscience* **36,** 12203–12216. ISSN: 0270-6474, 1529-2401. pmid: 27903729 (Nov. 2016).

51. Kühn, N. K. & Gollisch, T. Activity Correlations between Direction-Selective Retinal Ganglion Cells Synergistically Enhance Motion Decoding from Complex Visual Scenes. *Neuron.* ISSN: 0896-6273 (Jan. 2019).

52. Kunst, M., Laurell, E., Mokayes, N., Kramer, A., Kubo, F., Fernandes, A. M., Förster, D., Dal Maschio, M. & Baier, H. A Cellular-Resolution Atlas of the Larval Zebrafish Brain. *Neuron* **103,** 21–38.e5. ISSN: 08966273 (July 2019).

53. Land, M. Eye Movements in Man and Other Animals. *Vision Research* **162,** 1–7. ISSN: 0042-6989 (Sept. 2019).

54. Lappe, M., Pekel, M. & Hoffmann, K.-P. Optokinetic Eye Movements Elicited by Radial Optic Flow in the Macaque Monkey. *Journal of Neurophysiology* **79,** 1461–1480. ISSN: 0022-3077 (Mar. 1998).

55. Leyden, C., Brysch, C. & Arrenberg, A. B. A Distributed Saccade-Associated Network Encodes High Velocity Conjugate and Monocular Eye Movements in the Zebrafish Hindbrain. *Scientific Reports* **11,** 12644. ISSN: 2045-2322 (June 2021).

56. Longuet-Higgins, H. C. & Prazdny, K. The Interpretation of a Moving Retinal Image. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **208,** 385–397 (July 1980).

57. Lowe, A. S., Nikolaou, N., Hunter, P. R., Thompson, I. D. & Meyer, M. P. A Systems-Based Dissection of Retinal Inputs to the Zebrafish Tectum Reveals Different Rules for Different Functional Classes during Development. *Journal of Neuroscience* **33,** 13946–13956. ISSN: 0270-6474, 1529-2401 (Aug. 2013).

58. Marachlian, E., Avitan, L., Goodhill, G. J. & Sumbre, G. Principles of Functional Circuit Connectivity: Insights From Spontaneous Activity in the Zebrafish Optic Tectum. *Frontiers in Neural Circuits* **12,** 46. ISSN: 1662-5110 (2018).

59. Maris, E. & Oostenveld, R. Nonparametric Statistical Testing of EEG- and MEG-data. *Journal of Neuroscience Methods* **164,** 177–190. ISSN: 0165-0270. pmid: 17517438 (Aug. 2007).

60. Marr, D. & Poggio, T. From Understanding Computation to Understanding Neural Circuitry (May 1976).

61. Marr, D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* ISBN: 0-7167-1567-8 (Henry Holt and Co., Inc., USA, 1982).

62. Matsuda, K. & Kubo, F. Circuit Organization Underlying Optic Flow Processing in Zebrafish. *Frontiers in Neural Circuits* **15,** 73. ISSN: 1662-5110 (2021).

63. Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G. & Burr, D. C. A Cortical Area That Responds Specifically to Optic Flow, Revealed by fMRI. *Nature Neuroscience* **3,** 1322–1328. ISSN: 1097-6256. pmid: 11100154 (Dec. 2000).

64. Naumann, E. A., Fitzgerald, J. E., Dunn, T. W., Rihel, J., Sompolinsky, H. & Engert, F. From Whole-Brain Data to Functional Circuit Models: The Zebrafish Optomotor Response. *Cell* **167,** 947–960.e20. ISSN: 00928674 (Nov. 2016).

65. Neuhauss, S. C. F., Biehlmaier, O., Seeliger, M. W., Das, T., Kohler, K., Harris, W. A. & Baier, H. Genetic Disorders of Vision Revealed by a Behavioral Screen of 400 Essential Loci in Zebrafish. *Journal of Neuroscience* **19,** 8603–8615. ISSN: 0270-6474, 1529-2401 (Oct. 1999).

66. Nichols, T. E. & Holmes, A. P. Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples. *Human Brain Mapping* **15,** 1–25. ISSN: 1065-9471. pmid: 11747097 (Jan. 2002).

67. Orger, M. B., Smear, M. C., Anstis, S. M. & Baier, H. Perception of Fourier and Non-Fourier Motion by Larval Zebrafish. *Nature Neuroscience* **3,** 1128. ISSN: 10976256 (Nov. 2000).

68. Paninski, L. Convergence Properties of Three Spike-Triggered Analysis Techniques. *Network (Bristol, England)* **14,** 437–464. ISSN: 0954-898X. pmid: 12938766 (Aug. 2003).

69. Pnevmatikakis, E. A. *et al.* Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* **89,** 285–299. ISSN: 0896-6273 (Jan. 2016).

70. Portugues, R. & Engert, F. The Neural Basis of Visual Behaviors in the Larval Zebrafish. *Current Opinion in Neurobiology. Motor Systems • Neurology of Behaviour* **19,** 644–647. ISSN: 0959-4388 (Dec. 2009).

71. Prazdny, K. Egomotion and Relative Depth Map from Optical Flow. *Biological Cybernetics* **36,** 87–102. ISSN: 0340-1200, 1432-0770 (1980).

72. Reichardt, W. & Rosenblith, W. A. Sensory Communication (1961).

73. Reichardt, W. & Poggio, T. Visual Control of Orientation Behaviour in the Fly: Part I. A Quantitative Analysis. *Quarterly Reviews of Biophysics* **9,** 311–375. ISSN: 1469-8994, 0033-5835 (Aug. 1976).

74. Rister, J., Pauls, D., Schnell, B., Ting, C.-Y., Lee, C.-H., Sinakevitch, I., Morante, J., Strausfeld, N. J., Ito, K. & Heisenberg, M. Dissection of the Peripheral Motion Channel in the Visual System of Drosophila Melanogaster. *Neuron* **56,** 155–170. ISSN: 0896-6273 (Oct. 2007).

75. Roeser, T. & Baier, H. Visuomotor Behaviors in Larval Zebrafish after GFP-guided Laser Ablation of the Optic Tectum. *Journal of Neuroscience* **23,** 3726–3734. pmid: 12736343 (2003).

76. Ros, I. G. & Biewener, A. A. Pigeons (C. Livia) Follow Their Head during Turning Flight: Head Stabilization Underlies the Visual Control of Flight. *Frontiers in Neuroscience* **11,** 655. ISSN: 1662-453X (2017).

77. Royden, C. S. & Conti, D. M. A Model Using MT-like Motion-Opponent Operators Explains an Illusory Transformation in the Optic Flow Field. *Vision Research* **43,** 2811–2826. ISSN: 0042-6989 (Dec. 2003).

78. Rust, N. C., Schwartz, O., Movshon, J. A. & Simoncelli, E. P. Spatiotemporal Elements of Macaque V1 Receptive Fields. *Neuron* **46,** 945–956. ISSN: 0896-6273 (June 2005).

79. Rust, N. C., Schwartz, O., Movshon, J. A. & Simoncelli, E. Spike-Triggered Characterization of Excitatory and Suppressive Stimulus Dimensions in Monkey V1. *Neurocomputing. Computational Neuroscience: Trends in Research 2004* **58–60,** 793–799. ISSN: 0925-2312 (June 2004).

80. Samengo, I. & Gollisch, T. Spike-Triggered Covariance: Geometric Proof, Symmetry Properties, and Extension beyond Gaussian Stimuli. *Journal of Computational Neuroscience* **34,** 137–161. ISSN: 0929-5313. pmid: 22798148 (Feb. 2013).

81. Sherrington, C. *The Integrative Action of the Nervous System* (CUP Archive, 1952).

82. Simpson, J. I., Leonard, C. S. & Soodak, R. E. The Accessory Optic System of Rabbit. II. Spatial Organization of Direction Selectivity. *Journal of Neurophysiology* **60,** 2055–2072. ISSN: 0022-3077 (Dec. 1988).

83. Sjulson, L. & Miesenböck, G. Optical Recording of Action Potentials and Other Discrete Physiological Events: A Perspective from Signal Detection Theory. *Physiology (Bethesda, Md.)* **22,** 47–55. ISSN: 1548-9213 (Feb. 2007).

84. Srinivasan, M. V., Jin, Z. F., Stange, G. & Ibbotson, M. R. 'Vector White Noise': A Technique for Mapping the Motion Receptive Fields of Direction-Selective Visual Neurons. *Biological Cybernetics* **68,** 199–207. ISSN: 1432-0770 (Jan. 1993).

85. Sunkara, A., DeAngelis, G. C. & Angelaki, D. E. Joint Representation of Translational and Rotational Components of Optic Flow in Parietal Cortex. *Proceedings of the National Academy of Sciences* **113,** 5077–5082. ISSN: 0027-8424, 1091-6490. pmid: 27095846 (May 2016).

86. Takemura, S.-y., Karuppudurai, T., Ting, C.-Y., Lu, Z., Lee, C.-H. & Meinertzhagen, I. A. Cholinergic Circuits Integrate Neighboring Visual Signals in a Drosophila Motion Detection Pathway. *Current Biology* **21,** 2077–2084. ISSN: 0960-9822 (Dec. 2011).

87. Tanaka, K. & Saito, H. Analysis of Motion of the Visual Field by Direction, Expansion/Contraction, and Rotation Cells Clustered in the Dorsal Part of the Medial Superior Temporal Area of the Macaque Monkey. *Journal of Neurophysiology* (Sept. 1989).

88. Taylor, G. K. & Krapp, H. G. in *Advances in Insect Physiology* (eds Casas, J. & Simpson, S. J.) 231–316 (Academic Press, Jan. 2007).

89. Theis, L., Berens, P., Froudarakis, E., Reimer, J., Román Rosón, M., Baden, T., Euler, T., Tolias, A. S. & Bethge, M. Benchmarking Spike Rate Inference in Population Calcium Imaging. *Neuron* **90,** 471–482. ISSN: 0896-6273 (May 2016).

90. Theunissen, F. E., David, S. V., Singh, N. C., Hsu, A., Vinje, W. E. & Gallant, J. L. Estimating Spatio-Temporal Receptive Fields of Auditory and Visual Neurons from Their Responses to Natural Stimuli. *Network: Computation in Neural Systems* **12,** 289–316. ISSN: 0954-898X. pmid: 11563531 (Jan. 2001).

91. Vintch, B., Movshon, J. A. & Simoncelli, E. P. A Convolutional Subunit Model for Neuronal Responses in Macaque V1. *The Journal of Neuroscience* **35,** 14829–14841. ISSN: 0270-6474. pmid: 26538653 (Nov. 2015).

92. Wang, K., Hinz, J., Haikala, V., Reiff, D. F. & Arrenberg, A. B. Selective Processing of All Rotational and Translational Optic Flow Directions in the Zebrafish Pretectum and Tectum. *BMC Biology* **17,** 29. ISSN: 1741-7007 (Mar. 2019).

93. Warren, W. H. & Hannon, D. J. Direction of Self-Motion Is Perceived from Optical Flow. *Nature* **336,** 162–163. ISSN: 0028-0836, 1476-4687 (Nov. 1988).

94. Warren, W. H., Morris, M. W. & Kalish, M. Perception of Translational Heading From Optical Flow, 15.

95. Wertz, A., Haag, J. & Borst, A. Local and Global Motion Preferences in Descending Neurons of the Fly. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology* **195,** 1107–1120. ISSN: 1432-1351. pmid: 19830435 (Dec. 2009).

96. Wilt, B. A., Fitzgerald, J. E. & Schnitzer, M. J. Photon Shot Noise Limits on Optical Detection of Neuronal Spikes and Estimation of Spike Timing. *Biophysical Journal* **104,** 51–62. ISSN: 0006-3495 (Jan. 2013).

97. Winship, I. R. & Wylie, D. R. W. Receptive-Field Structure of Optic Flow Responsive Purkinje Cells in the Vestibulocerebellum of Pigeons. *Visual Neuroscience* **23,** 115–126. ISSN: 1469-8714, 0952-5238 (Jan. 2006).

98. Wu, Y., Dal Maschio, M., Kubo, F. & Baier, H. An Optical Illusion Pinpoints an Essential Circuit Node for Global Motion Processing. *Neuron* **108,** 722–734.e5. ISSN: 1097-4199 (Nov. 2020).

99. Wylie, D. R. W. Binocular Neurons in the Nucleus Lentiformis Mesencephali in Pigeons: Responses to Translational and Rotational Optic Flowfields. *Neuroscience Letters* **291,** 9–12. ISSN: 0304-3940 (Sept. 2000).

100. Wylie, D. R. W. & Frost, B. J. Responses of Neurons in the Nucleus of the Basal Optic Root to Translational and Rotational Flowfields. *Journal of Neurophysiology* **81,** 267–276. ISSN: 0022-3077 (Jan. 1999).

101. Xiao, Q. & Frost, B. J. Motion Parallax Processing in Pigeon (Columba Livia) Pretectal Neurons. *European Journal of Neuroscience* **37,** 1103–1111. ISSN: 1460-9568 (Apr. 2013).

102. Yildizoglu, T., Riegler, C., Fitzgerald, J. E. & Portugues, R. A Neural Representation of Naturalistic Motion-Guided Behavior in the Zebrafish Brain. *Current Biology* **30,** 2321–2333.e6. ISSN: 09609822 (June 2020).

103. Zanker, J. M., Srinivasan, M. V. & Egelhaaf, M. Speed Tuning in Elementary Motion Detectors of the Correlation Type. *Biological Cybernetics* **80,** 109–116. ISSN: 0340-1200, 1432-0770 (Feb. 1999).

104. Zhao, X., Chen, H., Liu, X. & Cang, J. Orientation-Selective Responses in the Mouse Lateral Geniculate Nucleus. *Journal of Neuroscience* **33,** 12751–12763. ISSN: 0270-6474, 1529-2401. pmid: 23904611 (July 2013).

105. van Santen, J. P. & Sperling, G. Elaborated Reichardt Detectors. *Journal of the Optical Society of America. A, Optics and Image Science* **2,** 300–321. ISSN: 0740-3232. pmid: 3973763 (Feb. 1985).

106. van den Berg, A. V. Robustness of Perception of Heading from Optic Flow. *Vision Research* **32,** 1285–1296. ISSN: 0042-6989 (July 1992).

Part V

ACKNOWLEDGEMENT

Tell him be brave
to dig deep,
his world awaits
in the deep down.

— Xiaojun Zhang,
(1925-2018)

5

ACKNOWLEDGEMENT

I would first like to acknowledge my grandma, Xiaojun Zhang, who ignited my interest in science, and encouraged me to pursue it in the past years and the recent dreams, and my parents Zhangzi Yang and Caien Zhang for their never stopped supports, believes and love. They are the source of my strength, without which I would not make half this far. I thank my wife who saves me from my pity and fear, and I wish we could stay young, go dancing.

I would like to especially thanks for my supervisor, Aristides Arrenberg, for not only his wise advices and intellectual support in these years, but giving me a high degree of freedom in research directions. None of these works can be done without his support and tolerance. I also want to thank Timm Schubert, who accurately pointed out my strength and weakness many years ago, and gave me many critical advices throughout my master and PhD. I would also thank my master supervisor, Thomas Euler, for teaching me how to analyze and visualize scientific data in a pretty and precise way. I must also express my gratefulness to Danielle Borg, my first mentor who encouraged me to stay in science and advised me to apply master programs in Germany when I was about to quit science because of the high education expenses in Australia.

I would like to thank Kun Wang and Julian Hinz, for a lot inspiring discussions with them. Many of their advices have turned out to be critical for my research project. I thank Ruoyu Huang for his contributions in building the behavioral setup and conducting many critical and difficult experiments. In the end, I'd like to thank my colleagues: Christian Brysch, Claire Leyden, Tim Hladnik and Giulia Soto, and my friends: Kaidi Shao, Yongrong Qiu, Zhijian Zhao, Ziwei Huang, Zuonan Cao, Yuchen Liu, Xingle Cai and May Li Silva-Prieto, I will miss the

life in the past 7 years in Tübingen with your company.

Part VI

LIST OF MANUSCRIPTS AND STATEMENT OF CONTRIBUTIONS

6

LIST OF MANUSCRIPTS AND STATEMENT OF CONTRIBUTIONS

6.1 HIGH-THROUGHPUT MOTION RECEPTIVE FIELD ESTIMATION

The article "High throughput, rapid receptive field estimation for global motion sensitive neurons using a contiguous motion noise stimulus." by **Yue Zhang** and Aristides B. Arrenberg, is published on *Journal of neuroscience methods* in October 2019.

Author contributions
Y.Z. and A.A conceived the project and wrote the manuscript. Y.Z. developed the CMN g stimulus and the two-step NCB statistic test. Y.Z. performed the *in vivo* calcium imaging in the optic tectum and pretectum of larval zebrafish and analyzed the data from the *in silico* and the *in vivo* experiments.

Appendix 7.1: We developed an efficient method for motion RF estimation which is compatible with high-throughput calcium imaging. The method contains two major components: the CMN stimulus and the two-step NCB test. The former is an advance feature noise stimulus which enables efficient motion RF estimation with reverse correlation. The latter serve as a quality control as well as a denoising algorithm that determine if the raw estimation contains any real RF structures and if so, where are they. The performance of this method has been evaluated with both *in silico* experiments and *in vivo* calcium imaging. And the results showed the RFs that represents the stimulus-response function of visual motion but not the other basic features in the stimulus can be efficiently and accurately identified by this method. This method makes systematic motion RF characterization with *in vivo* calcium imaging possible, which is critical for the understanding of optic flow processing at the algorithmic level.

6.2 PARALLEL CHANNELS FOR MOTION FEATURE EXTRACTION

The article "Parallel Channels for Motion Feature Extraction in the Pretectum and Tectum of Larval Zebrafish" by Kun Wang, Julian Hinz, **Yue Zhang**, Tod R. Thiele and Aristides B. Arrenberg, is published on *Cell Report* in January 2020.

Author contributions
K.W. performed the calcium imaging experiments on pretectal and tectal somatic responses as well as the behavioral experiment. K.W. and J.H. analyzed the calcium imaging data. Y.Z. contributed to the tail tracking algorithm for behavioral data analysis, and K.W. analyzed the behavioral data. A.B.A., K.W., J.H., and T.R.T. conceived the experiments and associated analysis protocols. K.W., J.H., and A.B.A. wrote the manuscript, with input from T.R.T.

Appendix 7.2:The brain regions optic tectum and pretectum are heavily involved in visual motion processing in larval zebrafish. In this paper, the RF of the tectal and pretectal neurons sensitive to horizontal motion in the right visual fields are characterized with the calcium responses to the moving grating stimulus with various size and presented at different locations. The RFs of the motion sensitive neurons in the optic tectum are biased towards the upper visual field. And they only respond to visual motion in a relatively small region but not the wide-field motion covering the same local region. While the pretectal motion-sensitive neurons tend to have large RFs covering lower visual field, which indicates their role in processing global motion. This is further supported by the behavioral experiment in which the OMR can be more frequently evoked by the forward translation-induced optic flow visible in the lower temporal visual field. These results indicate the pretectal and tectal motion sensitive neurons extract different visual motion features in parallel and form differential representations of optic flow.

6.3 A ROBUST RECEPTIVE FIELD CODE FOR OPTIC FLOW DECOMPOSITION

The article "A robust receptive field code for optic flow detection and decomposition during self-motion." by **Yue Zhang**, Ruoyu Huang, Wiebke Nörenberg and Aristides B. Arrenberg, is published on *bioRxiv* in October 2021.

Author contributions
Y.Z. and A.A conceived the project and wrote the manuscript. Y.Z. R.H. and W.N. performed the *in vivo* calcium imaging in the optic tectum and pretectum of larval zebrafish. Y.Z. and R.H. performed the behavioral experiment that evaluate the OKR and OMR to self-motion induced optic flow. R.H. performed the spinal cord injection for retrograde labelling of the nucleus of the medial longitudinal fasciculus (nMLF). Y.Z. developed the CMN and self-motion induced optic flow stimulus for all neurophysiological and behavioral experiment as well as the recording software for behavioral experiments and analyzed the data from all experiments.

Appendix 7.3:Through the systematic RF characterization of optic tectum and pretectal area in larval zebrafish, it has been revealed that the neurons involved in global motion processing implement the matched filter algorithm for optic flow detection with the RF structures similar to specific self-motion induced optic flow. The neural response to self-motion induced optic flow as well as the related tuning properties can be largely explained by these linear filter RFs. The neurons sensitive to translation-induced optic flow are topographical arranged in the pretectal AMC according to the translation direction. The population code of these neurons is sufficient and robust for representing self-translation direction. We further demonstrated that the algorithm implemented by these neurons satisfied the computational principle of decomposability with a behavioral experiment in which larval zebrafish responded to the rotation-induced optic flow component with OKR and translational component with OMR when the two components are simultaneously presented in the optic flow. This decomposition also suggests that the OKR and OMR, or at least one of them, are not responding to local motion but the self-motion information extracted from the global optic flow. These results deepen the understanding of optic flow processing in vertebrate species at the computational,

algorithmic and implementational levels.

Part VII

APPENDIX

7

APPENDIX

7.1 HIGH THROUGHPUT, RAPID RECEPTIVE FIELD ESTIMATION FOR GLOBAL MOTION SENSITIVE NEURONS USING A CONTIGUOUS MOTION NOISE STIMULUS

Journal of Neuroscience Methods 326 (2019) 108366

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

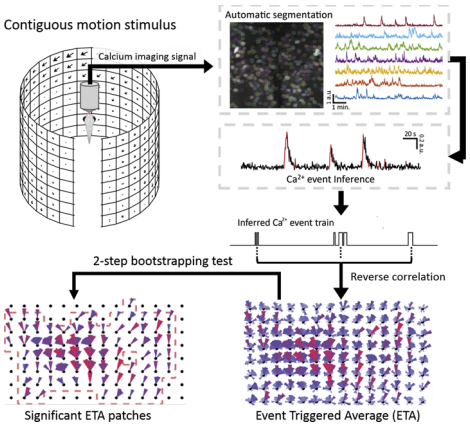
journal homepage: www.elsevier.com/locate/jneumeth

High throughput, rapid receptive field estimation for global motion sensitive neurons using a contiguous motion noise stimulus

Yue Zhang[a,b], Aristides B. Arrenberg[a,*]

[a] Werner Reichardt Centre for Integrative Neuroscience, Institute of Neurobiology, University of Tübingen, D-72076, Tübingen, Germany
[b] Graduate Training Centre for Neuroscience, University of Tübingen, D-72076, Tübingen, Germany

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Receptive field
Reverse correlation
Motion vision
Direction-selective
Noise stimulus
Calcium imaging
Optic flow
Pretectum

ABSTRACT

Background: The systematic characterization of receptive fields (RF) is essential for understanding visual motion processing. The performance of RF estimation depends on the employed stimuli, the complexity of the encoded features, and the quality of the activity readout. Calcium imaging is an attractive readout method for high-throughput neuronal activity recordings. However, calcium recordings are oftentimes noisy and of low temporal resolution. The RF estimation of neurons sensitive to global motion is particularly challenging due to their potentially complex combination of preferred directions across visual field positions.

New method: Here, we present a novel noise stimulus, which is enriched with spatiotemporally contiguous motion and thus triggers robust calcium responses. We combined this contiguous motion noise (CMN) stimulus with reverse correlation followed by a two-step nonparametric cluster-based bootstrapping test for efficient and reliable RF estimation.

Results: The *in silico* evaluation of our approach showed that RF centre positions and preferred directions are reliably detected in most of the simulated neurons. Suppressive RF components were detected in 40% of the simulated neurons. We successfully applied our approach to estimate the RFs of 163 motion-sensitive neurons *in vivo* within 40 min in the pretectum of zebrafish. Many *in vivo* neurons were sensitive to elaborate directional flow fields in their RFs.

Comparison with existing methods: Our approach outperforms white noise methods and others due to the

* Corresponding author.
E-mail address: aristides.arrenberg@uni-tuebingen.de (A.B. Arrenberg).

https://doi.org/10.1016/j.jneumeth.2019.108366
Received 16 May 2019; Received in revised form 9 July 2019; Accepted 14 July 2019
Available online 26 July 2019

optimized motion stimulus statistics and ascertainable fine RF structures.
Conclusions: The CMN method enables efficient, non-biased RF estimation and will benefit systematic high-throughput investigations of RFs using calcium imaging.

1. Introduction

One primary task of visual information processing is to extract low-dimensional features from the high-dimensional light patterns present in the environment. The concept "receptive field" (RF) was proposed to quantitatively describe the spatiotemporal stimulus structures that modulate neural activity in a single neuron (Hubel and Wiesel, 1962). For instance, the RF structures of orientation-selective neurons are similar to the stimulus patterns that evoke firing (Bonin et al., 2011; Hubel and Wiesel, 1962), and the RF structures of local motion sensitive neurons are consistent with spatiotemporal structures of the locally moving patterns these neurons respond to (Adelson and Bergen, 1985). This quantitative description is critical for understanding mechanisms of the visual system, since a plethora of different RFs exist within and across visual brain areas, which together form the basis of feature extraction. The measurement of receptive fields critically depends on the stimulus statistics, the readout of cellular activity, and the algorithm used to infer the receptive field based on stimulus and measured activity. The stimulus statistics affect the RF estimation in several ways: i) some stimuli are more effective at driving responses than others given the limited experimental time, ii) the neural responses to more complex stimuli can provide more information about the corresponding RF, iii) the biases in the stimulus statistics may be inherited in the RF estimation result.

From studies in Drosophila, zebrafish, pigeon and human, it is known that the processing of global motion is critical for the visual control of locomotion behaviour (Frost et al., 1994; Kubo et al., 2014; Warren et al., 2001; Wertz et al., 2009a). In previous studies, it has been found that the pretectal neurons in the larval zebrafish brain are involved in processing horizontal global motion (Kubo et al., 2014; Portugues et al., 2014) to mediate optokinetic and optomotor stabilization behaviours. While the RF structures of many neuron types have been measured, *e.g.* those of mammalian retinal ganglion cells and V1 simple cells (Hubel and Wiesel, 1962; Rust et al., 2005), the RF structures of large-field motion-sensitive neurons of the pretectum and accessory optic system are, however, mostly unknown. The lack of an efficient method for estimating relatively fine structure of large-size motion receptive fields has impeded their systematic and comprehensive characterization. An improved motion RF estimation method may greatly benefit our understanding of motion processing mechanisms in the brain, including visuomotor transformations underlying locomotor and stabilization behaviours.

Arguably the most common receptive field estimation technique is reverse correlation, which has frequently been employed to estimate spatiotemporal receptive fields of direction-selective visual neurons (DeAngelis et al., 1993; Eckhorn et al., 1993; Rust et al., 2005; Wang and Yao, 2011). Provided that neurons respond to spatiotemporal, random noise stimuli, the linear receptive field structures of most visual neurons in the early visual system can be estimated by computing the averaged patterns of the spike-associated stimuli (Boer and Kuyper, 1968; Chichilnisky, 2001; Marmarelis and Naka, 1972). The efficiency of this method is highly dependent on the frequency of neural activity events during stimulation and on the information density of the stimulus, which both together determine the acquisition rate of RF-relevant information during noise stimulation. Naturally, this information acquisition rate is affected by the relevance of the noise stimuli to the features the measured neurons prefer. For example, white noise stimuli are not efficient for estimating receptive fields of neurons preferring low spatial frequency features. Much of the recording time would be "wasted" on irrelevant high frequency stimulus features and the neurons would mostly be inactive during the recording. Thus, for a time-efficient receptive field estimation, it is essential to constrain the noise stimuli to match the preferred features of the neurons in question.

A common constraining approach is to manually decompose the possible feature-relevant sensory space to a set of simple stimuli defined by a small number of parameters of interest. Based on the neuronal responses to these stimuli, the RF properties can be characterized regarding these parameters of interest. For example, the preferred directions of direction-selective retinal ganglion cells in rabbit can be estimated from their responses to light spots moving in 8 different directions (Barlow et al., 1964). In a similar way, motion stimuli can be presented in defined local regions to estimate direction selectivity profiles across visual field locations for individual neurons. Using this approach, receptive field structures of global motion sensitive neurons in Drosophila have successfully been estimated (Krapp et al., 1998; Krapp and Hengstenberg, 1996; Wertz et al., 2009b).

Since the static features of motion stimuli are less relevant for estimating the direction-selective receptive field components, the expression of stimuli by two-dimensional motion vectors within the visual field represents a valid simplification and reduction of parameter space. Correspondingly, vector noise stimuli randomly sampled from motion vector space have been created for motion RF estimation previously (Borghuis et al., 2003; Srinivasan et al., 1993).

However, these methods heavily rely on the high sensitivity and temporal resolution of electrophysiological neural activity recordings. Electrophysiological recordings are oftentimes time-consuming or limited to recording one neuron at a time, which makes the approach laborious for the systematic investigation of vertebrate brain areas containing high numbers of neurons. Calcium imaging is a promising method for functionally characterizing such large numbers of neurons within a given brain area in parallel (Froudarakis et al., 2014; see Grienberger and Konnerth, 2012 for review). Recent developments in imaging techniques (Ahrens et al., 2013; Akerboom et al., 2012; Chen et al., 2013), including genetically encoded calcium indicators and improved sample preparation techniques, make two-photon calcium imaging a sharp tool for studying vision processing *in vivo* (Ahrens et al., 2013; Baden et al., 2013; Bonin et al., 2011; Froudarakis et al., 2014; Hunter et al., 2013). However, the relatively poor sensitivity of calcium indicators and the low temporal resolution of calcium imaging limit its power in detecting and/or resolving individual action potential spikes, which would be needed for the receptive field estimation methods described above.

Here, we present a novel approach for the efficient estimation of receptive fields of global motion sensitive neurons using *in vivo* calcium imaging. We employed a contiguous motion vector noise stimulus, which can efficiently trigger strong, robust neuronal calcium responses in global motion sensitive neurons. With this stimulus and a dedicated bootstrapping analysis method, we successfully estimated fine-scale motion receptive field structures using noisy two-photon calcium imaging at low temporal resolution (2 Hz) in the global motion sensitive pretectum of larval zebrafish.

2. Materials and methods

2.1. Ethical approval

All animal experiments were licensed by the local authorities (Regierungspräsidium Tübingen) in accordance with German federal law and Baden-Württemberg state law.

2.2. Visual stimulation setup

Visual stimuli were displayed on a custom-built cylindrical green LED arena covering 336°-by-80° of the visual field (-168° to 168° in azimuth; -40° to 40° in elevation). The arena consisted of 224 square LED tiles (8 rows × 14 columns × 2 arena halves); each tile had 8 × 8 evenly distributed LEDs emitting at 570 nm (Kingbright TA08-81CGKWA); On average, each LED covered ˜1.5 horizontal visual degrees of visual field. The vertical visual angle each LED covered varied slightly with its latitude. Because the LED was programmed to only display two luminance levels (on/off), all stimuli displayed in the arena had 100% contrast. All LEDs were covered with a high-pass filter foil (LEE no. 779, article 595-1700-7790, castinfo.de, Hagen, Germany) to reduce the light interference from the LEDs at the fluorescence emission wavelength. An additional diffusion filter foil (LEE no. 252, article 595-1780-2520) was also placed in front of the LED tiles to homogenize the LED light. Custom software developed in previous studies (Joesch et al., 2008; Reiser and Dickinson, 2008) was used for uploading, displaying and controlling the visual stimulus. To minimize the interference of LED light with the calcium imaging, the LEDs were flickering in a duty cycle such that they were switched on only during the fly-back time of the scanning mirrors in-between line scans. Before the experiment started, the spherical fish container (a glass bulb) was fixed to a metal holder surrounded by the arena. Tilts of the fish head in the pitch, yaw and roll axis were corrected under the wide field microscope.

2.3. Contiguous motion noise

The motion-rich noise stimulus in this study shared the same principle as the vector white noise introduced by Srinivasan and colleagues (1993), which constrains the stimulus space to contain only first-order motion. Instead of white noise, we used correlated noise to present locally coherent motion stimuli, which are common in natural ego-motion scenes and enable more efficient and robust receptive field estimation of global motion sensitive neurons using reverse correlation and calcium imaging. We named our stimulus "Contiguous Motion Noise Stimulus" (CMN stimulus, Fig. 1, Supplementary Video 1). To construct the correlated vector noise, we first randomly sampled a number of motion vectors whose norm and phase angle defined moving speed and direction, respectively, from a circular uniform distribution:

$$P(\Theta = \theta) = \frac{1}{2\pi}, \ where\ \theta \in [-\pi, \pi]$$

The Θ corresponded to the phase angle of motion vectors. The norm of all motion vectors was set to the same arbitrary value of 1. Unifying the motion vector norm at this stage helped to control the distribution of speeds in the following steps (see further explanations below), and the value was an arbitrary number since the vector norms will be normalized in a later step.

These motion vectors were organized into a three-dimensional motion vector matrix with the configuration: 16 rows × 38 columns × 12,020 frames (elevation × azimuth × time). The motion vector matrix was convolved with a modified multivariate Gaussian kernel to enforce the local coherences of the motion vectors:

$$f(Z) = exp(-\frac{1}{2}Z^T\Sigma^{-1}Z)$$

where:

$$Z = \begin{bmatrix} x \\ y \\ t \end{bmatrix} and\ \Sigma = \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_y^2 & 0 \\ 0 & 0 & \sigma_t^2 \end{bmatrix}$$

The spatial and temporal standard deviation of the Gaussian kernel, denoted as σ_s and σ_t respectively ($\sigma_s = \sigma_x = \sigma_y$), were set to π/2 vector distance (≈18.8 visual degrees) and 5 frames (≈ 0.33 s, details of the unit conversion are provided in the later sections), respectively. To minimize the edge effect in the convolution, an 8 × 28 × 12,000 matrix was cropped out from the centre of the convolved matrix and used as the motion vector matrix for later computation. The norm of the motion vectors in the convolved matrix was normalized to the range [0, k], with k corresponding to the maximum speed in the stimulus. In our CMN stimulus, k was set to 67 visual degrees/second.

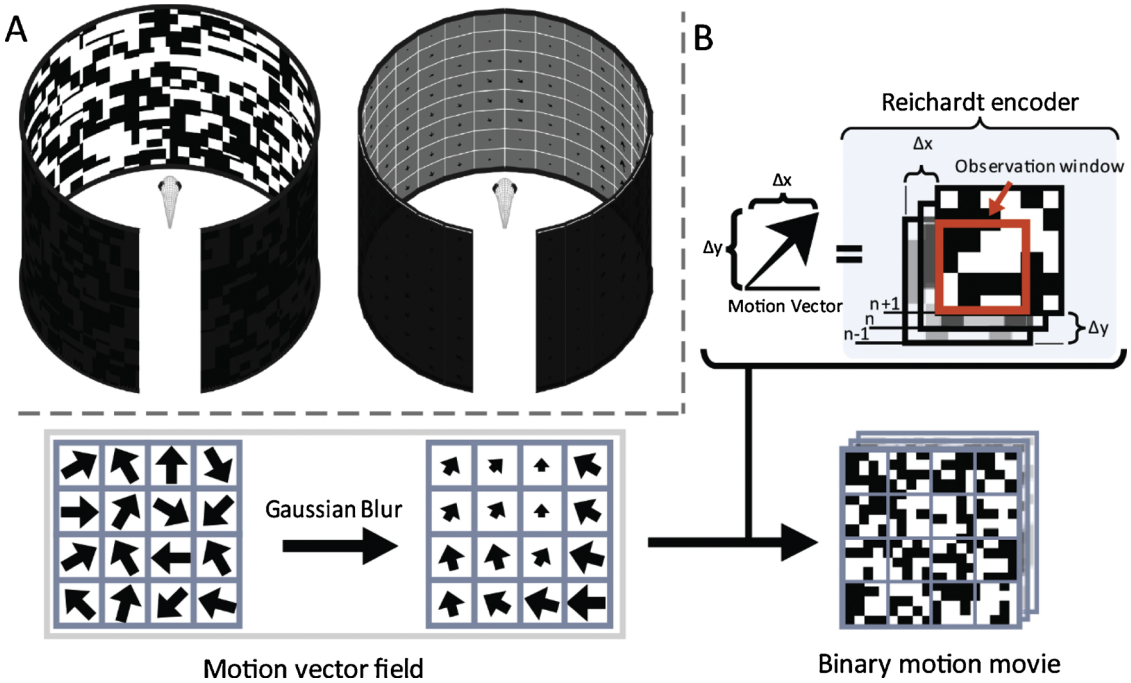


Fig. 1. Contiguous Motion Noise (CMN) stimulus design. (A) A single frame of the binary Contiguous Motion Noise (CMN) movie displayed on a cylindrical LED arena (left). The moving direction and speed of motion displayed on the LED arena are illustrated by the phase angle and length of the motion vectors in the right plot, respectively. (B) Illustration of the procedure for stimulus generation. A random motion vector field is convolved with a multivariate Gaussian function. The resulting contiguous motion vector field is transformed into a patterned binary motion movie making use of a Reichardt encoder.

The correlations between positions and frames of the motion matrix were calculated and shown as auto correlograms (Fig. 2C). To quantitatively describe the local coherence intensity, the radius of the strongly correlated region where the correlations between the centre vector and all the vectors in this region were higher than a threshold *ct*, was referred to as the contiguity radius (CR) of the spatial and temporal domain of the stimulus. Since the coherence in the stimulus was introduced by a single source (the multivariate Gaussian kernel), the spatial and temporal correlation can be well approximated by the autocorrelation of the Gaussian kernel in the corresponding dimension. The CR for the dimension can then be calculated as:

$$\begin{aligned}(f \star f)(\mathbf{CR}) &= ct \\ \exp(-\frac{\mathbf{CR}^2}{4\sigma^2}) &= ct \\ \mathbf{CR} &= 2\sigma\sqrt{-\ln(ct)}\end{aligned}$$

The σ was the spatial or temporal standard deviation of the multivariate Gaussian kernel. The comparison between the approximated and the exact CR is shown in Fig. 2C. In our stimulus, the threshold *ct* was set to 0.1. The spatial CR was $2\sigma_s\sqrt{-\ln(0.1)} = 4.77$ vector distances ≈ 57 visual degrees; the temporal CR was $2\sigma_t\sqrt{-\ln(0.1)} \approx 15$ frames = 1 s.

Because the convolved motion vector norm can be considered as a weighted mean of multiple random vectors sampled from the circular uniform distribution, the frequency distribution of convolved motion vector norms can be modelled by a Rayleigh distribution s(x):

$$s(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}},\ x \in [0,\ k],$$

where $\sigma \approx 0.1852k$, when k = 67 visual degrees/second, σ = 12.35 visual degrees/second. The cumulative distribution function S(x) was:

$$S(x) = 1 - e^{-\frac{x^2}{2\sigma^2}}$$

The s(x) and S(x) provided analytical descriptions of the motion speed distribution in the CMN stimulus. These descriptions can be used to quantify the statistics of motion speed presented in the CMN stimulus as follows: Let S(x) = y, then:

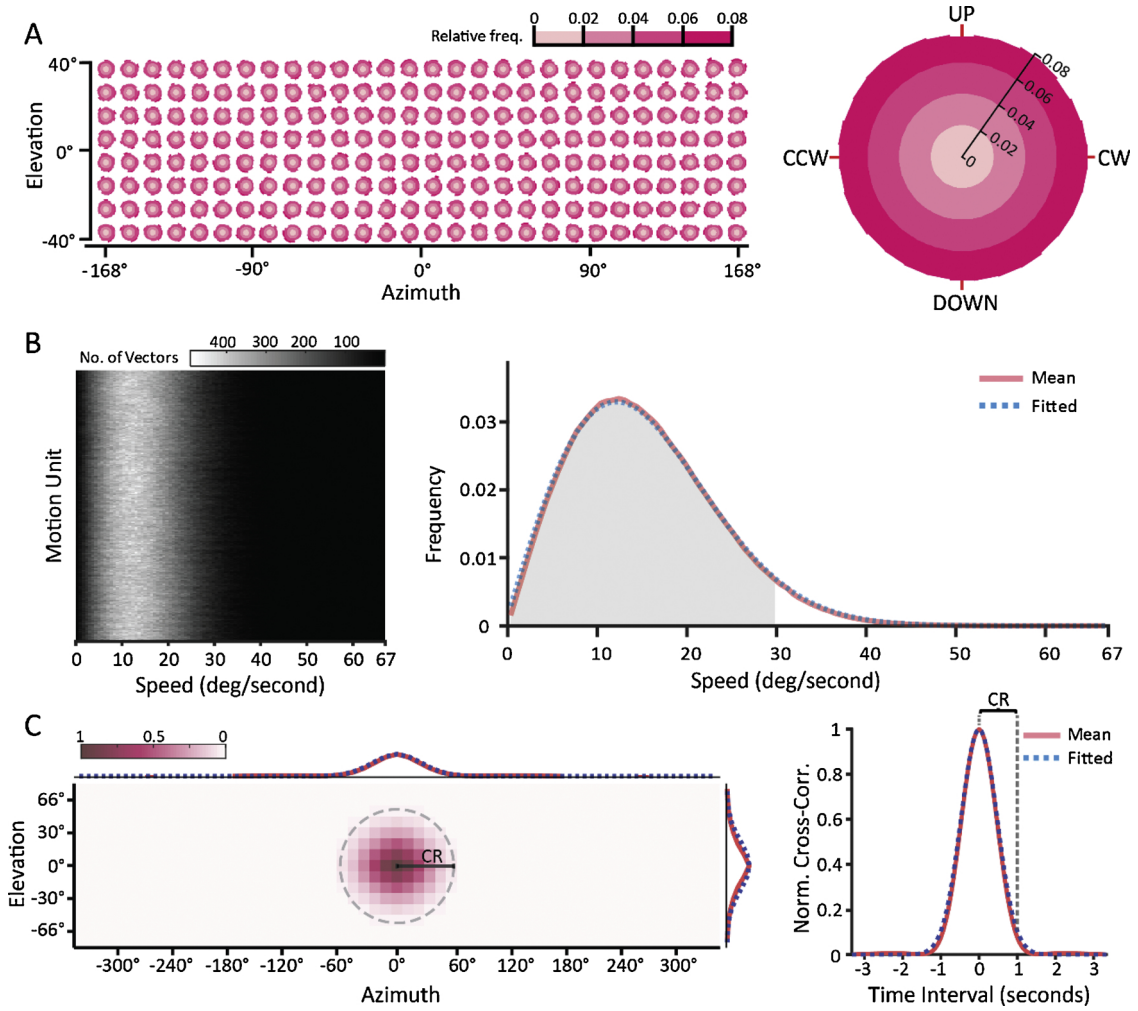


Fig. 2. Stimulus statistics of the Contiguous Motion Noise (CMN) stimulus used in this study. **(A)** Polar histograms of motion direction distributions for each spatial location (left) and all locations in in the motion vector matrix (right) for constructing the CMN stimulus movie (12,000 frames ≈ 13 min). The relative frequency of each motion direction is represented by the colour and the bar height. **(B)** The motion speed distribution for each spatial location (left) and the averaged speed distribution for all locations (right, red curve) in the CMN movie. The blue curve corresponds to the modelled speed distribution using the Rayleigh distribution function. **(C)** The spatial (left) and temporal (right) cross-correlogram of motion vectors. The elevation- and azimuth-averaged cross-correlograms (red curves) for each spatial dimension are shown above and next to the left plot, respectively. The blue dashed curves correspond to the cross-correlation modelled with the autocorrelation of the multivariate Gaussian kernel used in the stimulus construction (see Materials and methods for details). The contiguous radii (CRs) for the spatial and temporal domain calculated based on these models are indicated. The goodness of fit of these models was 0.99 (elevation-averaged), 0.90 (azimuth-averaged) and 0.99 (azimuth/elevation-averaged).

$$y = 1 - e^{-\frac{x^2}{2\sigma^2}}$$

$$x = \sqrt{-2\sigma^2 \ln(1-y)}$$

When x equalled 0.4534k and 1k, S(x) equalled to 0.95 and 1–4.67 × 10^{-7} ≈ 1 respectively. The optokinetic responses (OKR) of larval zebrafish can be readily triggered with grating stimuli moving at speeds around 20–30 visual degrees/second (Qian et al., 2005). Therefore, we set k to 67 visual degrees/second, so that 95% of the CMN stimulus motion vector samples will have speeds in the range of (0-0.4534)k, which equals to 0–30.38 visual degrees/second.

2.4. Reichardt encoder

The contiguous motion vector matrix constructed in the previous section needs to be converted to a movie in a way that the motion information neurons decode from the movie should be the same as motion information in the motion vector matrix. In a wide range of species, the motion detection is dominated by processing first-order motion cues (for zebrafish, see Orger et al., 2000). Several linear operators, including the elaborated Reichardt detectors (ERD) model, were developed to model the first order motion decoding in the visual system (Adelson and Bergen, 1985; Reichardt and Rosenblith, 1961; van Santen and Sperling, 1985). The ERD and other operators share a similar principle, which involves converting the auto-correlated spatiotemporal structure in the visual input to abstract first-order motion information. It is thus possible to construct an inverse operator of the ERD which converts the abstract motion information back to spatiotemporal structures that can be detected by first-order motion sensitive neurons. We used the term "Reichardt encoder" to refer to this first order motion encoder.

The Reichardt encoder we constructed was similar to the motion reverse correlation stimulus used in a previous study (Borghuis et al., 2003). Both of them present motion vectors with sequences of displacements of noise pattern; the movement each motion vector encoded was approximated by the displacement distance between the frames before and after the current frame divided by the time interval for each moment (Fig. 1B). The luminance of each pixel on the noise pattern image was assigned to a random binary value (0 or 1). To trigger motion perception properly in larval zebrafish, the maximum spatial frequency of the binary noise pattern was set to 0.1 cycles per visual degree by resizing patterns with a scale factor of 3 (Rinner et al., 2005). The resized noise image contained 3000 × 3000 pixels.

For each motion vector in the CMN motion vector matrix, the patterns in an 8 × 8 pixels observation window on a copy of the resized noise image was displayed on an 8 × 8 LED array (covering ~12 visual degrees in azimuth). In the first frame of the CMN stimulus, the observation window of each motion vector was positioned at a random place on the copy of the noise image. The displacement sequences for each copy of noise image were independent from each other. To solve the boundary problem (occurring when the boundary of the copied noise image ran into the observation window) we used circular boundary conditions, *i.e.* the boundaries of each copy of the noise image were connected with the boundaries at the opposite side forming a torus. The displacement distances of the resized pattern were rounded to multiples of 1.2 visual degrees (the resolution of the LED arena).

Each 8 × 8 LED array thus corresponded to one motion vector of the CMN stimulus and the motion vector matrix was converted to a 64 × 224 × 12,000 (pixel rows – columns – frames) binary noise movie displayed on the 64 × 224 pixel LED arena (corresponding to 8 × 28 motion vectors). Since copies of the noise image were displaced independently of each other in adjacent motion vector positions, it was unavoidable that the CMN movie contained different and varying luminance structures at the border between two 8 × 8 LED arrays, resulting (among other effects) in a “checkerboard” appearance of the time-averaged CMN movie (Figure S1). The stimulus movie was displayed at 15 frames per second (15 Hz), resulting in a stimulus duration of 13.3 min. The motion encoded in the first and the last columns of the motion vector matrix was displayed at the -168° to -156° region and the 156° to 168° region (in azimuth) on the cylindrical LED arena, respectively. The motion in the top and bottom rows of the motion vector matrix was displayed at the 32°to 40° region and the -32° to -40° region (in elevation).

2.5. Linear-nonlinear-Poisson model for the in silico experiment

We used the linear-nonlinear-Poisson (LNP) model to simulate neuronal responses to the CMN stimulus in the *in silico* experiments in a similar way as described in a previous study (Chichilnisky, 2001). The linear receptive field of a simulated neuron can be expressed as a matrix $\mathbf{w}$; each row of the matrix is a two-element vector $\mathbf{w}_i$. The norm and the angle of $\mathbf{w}_i$ indicated its weight and preferred direction for the spatial location i in the visual field. The norm of w_i should be in the range of [0,1]. Similarly, each frame of the motion vector matrix of the CMN stimulus can be transformed into a matrix $\mathbf{s}$. Its row vector $\mathbf{s}_i$ describes the movement at the spatial position i in the CMN stimulus. The norm of all motion vectors in the CMN stimulus was normalized to [0,1] by min-max rescaling here. The output of the linear receptive field, $f(\mathbf{s})$ was calculated as:

$$f(\boldsymbol{s}) = \sum \boldsymbol{w}_i \cdot \boldsymbol{s}_i^T \quad (1)$$

For the visual motion neurons with suppressive receptive field components, the output of the linear-nonlinear part was modified as follows:

$$f(\boldsymbol{s}) = max\left(\sum \boldsymbol{w}_i^+ \cdot \boldsymbol{s}_i^T, 0\right) - max\left(\sum \boldsymbol{w}_i^- \cdot \boldsymbol{s}_i^T, 0\right)$$

The $\mathbf{w}^+$ and $\mathbf{w}^-$ in the equations were the linear excitatory and suppressive receptive field structures. The positive linear rectifiers in $f(\mathbf{s})$ constrain the $\mathbf{w}^+$ and $\mathbf{w}^-$ to only produce excitatory and suppressive output respectively. Since the receptive field structure of the pretectal neurons had not been revealed in any previous study, here we simply simulated the excitatory and suppressive receptive field structures as containing Gaussian shapes, and all units in the same simulated receptive field component shared the same direction preferences. The linear filter with two excitatory components was used for simulating the binocular motion RFs; the preferred directions of two components were set to two random directions. The size of each of these Gaussian RF components was equal to $\sqrt{-\ln 0.1}\,\sigma = 1.52\,\sigma$ vector distance ≈ 18.2 σ visual degrees. We created simulated neurons with σ = 1.5, 4.5, 9; the receptive field of these simulated neurons covered approximately 27, 82 and 164 visual degrees respectively.

In the absence of normalization, the magnitude of output of large linear receptive fields can be larger than the output of receptive fields with smaller size, which would have resulted in higher firing rates in simulated neurons with larger receptive fields. To account for this problem in our simulation, the linear system output $f(\mathbf{s})$ was divided by $f(\mathbf{w})$. The normalized outputs $f_{norm}(s)$ were then rectified by the nonlinear unit $g(\mathbf{s}) = max(f_{norm}(s), c)$, c is the normalized rectifying threshold; in our simulated neuron model, c = 0.1, so g(**s**) ranged between 0.1 and 1. The spike train was generated from this normalized output by a Poisson spike generator.

To mimic the slow kinetics of calcium indicator signals and the noisy environment in the *in vivo* calcium imaging recordings, the spike train was first convolved with a calcium decaying kernel which was an exponential decay function with a mean lifetime of 3 s. Gaussian white noise was added before and after the convolution to mimic different source of noises in the calcium recordings. The span of the Gaussian noise distribution was set to 20% of the span of signal distribution. The convolved signal was down-sampled from the stimulus displaying rate of 15 Hz to the calcium image acquisition rate of 2 Hz. This down-

sampled signal trace represented the final output of the simulated neurons in the *in silico* experiment and was used for all following data analysis.

2.6. Animal preparation and in vivo two-photon calcium imaging

Zebrafish (*Danio rerio*) were maintained on a 14 h light–dark cycle at 28.5 °C and experiments were performed at room temperature. The transgenic zebrafish line *Tg(elavl3:nls-GCaMP6s)mpn400* was used in this study (Dal Maschio et al., 2017). Larval zebrafish which had strong calcium indicator expression levels and good optokinetic responses were selected from all hatched fish at 4-day post-fertilization (dpf), resulting in selection of the 10–20% best fish for experiments.

The *in vivo* two-photon calcium imaging experiment was performed on the selected larval zebrafish at 6 dpf. On the experiment day, the screened fish were immobilized by mounting them in 1.6% low melting-temperature agarose (wt/vol, E3 medium) at the tip of a transparent plastic triangle stage (tip angle < 45%). Then the triangle stage with the mounted fish was transferred into a spherical glass container filled with E3 medium. The containers were placed to the centre of the cylindrical LED arena. The body position and orientation of the mounted fish was adjusted later, so it was located in the centre of the LED arena with the dorsal side facing up (roll correction) and the head pointing to 0° elevation (pitch correction) and 0° azimuth (yaw correction).

The calcium activities of neurons in 3 horizontal planes at different depths of the pretectal area in a larval zebrafish were recorded with a movable objective microscope two-photon setup (Sutter Instruments, Novato, California, USA; Euler et al., 2009) coupled to a Coherent Vision-S Ti-Sa laser. The emitted fluorescence light of the calcium indicator protein, GCaMP6s, triggered by the excitation laser beam at 920 nm (prepulse compensation: 9756 fs^2) was collected with a 20x/1.0 Zeiss objective at the wavelength 500–550 nm. The image time series were recorded using MScan software (Sutter instruments). The image acquisition rate was 2 frames per second at 512 × 512 $pixel^2$ and a magnification of 2 (0.44 × 0.44 μm^2 per pixel).

2.7. Preprocessing of in vivo and in silico data

The recorded *in vivo* calcium imaging data was registered *via* cross-correlation of the frames to get rid of motion artefacts in the XY plane (rigid XY translation). All visible neuronal soma bodies in the time-averaged image of the registered calcium image time series were automatically selected as regions of interests (ROIs) by a marker-controlled watershed algorithm custom written using MATLAB. In total, 1896 ROIs were selected in the three transverse planes recorded. The calcium time series was computed for each ROI as the sum of all pixel values in each frame within the ROI. The calcium time series were each normalized to the range of [0, 1].

Due to a number of reasons (slow kinetics of the calcium indicator, low affinity of the calcium indicator, low image acquisition rate of our two-photon microscope, recording noise), calcium activities corresponding to single action potentials should be hard to detect. We therefore used the abrupt, high amplitude increases in the calcium traces (which we call events) as indicator for neural activities. These events should correspond to sudden, significant increases of the fluorescence signal related to the increases of neural firing rate, which were more robust to noise and more suitable for the low temporal resolution of the recording system than single action potentials. Note that we exclusively use the increases of activity and ignored the temporal profile of the calcium signal decay after each event. These calcium events were automatically identified by a Gaussian-mixture-model (GMM) based clustering algorithm. We assumed the noise in calcium recordings was Gaussian distributed according to the central limit theorem. Any significant increases of neuronal firing rate would drive the calcium signals away from the original Gaussian noise distribution. With this assumption, we assigned all data points in each calcium trace into one of the two Gaussian distributions by performing clustering with a GMM model. The outliers can be detected as the ones that are less likely to be drawn from the baseline Gaussian distribution and therefore more likely to correspond to the other (signal) distribution.

The GMM algorithm did not consider the temporal structure of signals, thus the detection can be strongly interfered by low frequency noise that caused baseline changes. To compensate for this, we applied the GMM clustering algorithm to all signals in a 200-frame moving window with the moving step size = 1 recording frame; each data point was thus assigned repeatedly for 200 times. The dominant assignments in these 200 repeats was used as the final assignment. The final output of the algorithm was a binary array (1 = calcium event detected; 0 = no event detected).

This calcium event detection algorithm was applied to both the simulated *in silico* and the recorded *in vivo* calcium traces. The accuracy of this event detection algorithm was tested in 50 *in silico* calcium traces. The inferred calcium event trains were up-sampled with the nearest-neighbour resampling method from the image acquisition rate (2 Hz) to the stimulus frame rate (15 Hz) to be aligned with the stimulus frames. On average, each simulated neuron had calcium events in 35.1 ± 4.5% of the stimulus frames (mean ± standard deviation, same in the other sentences and paragraphs). The detection accuracy calculated as the number of correctly assigned frames divided by the total number of assigned frames was 77.80 ± 1.84% for the *in silico* data. The false positive rate was 20.77 ± 2.93%. The high false positive rate was mainly caused by the nearest-neighbour up-sampling of the inferred calcium event train. If we calculate the false positive rate in the stimulus frames down-sampled to the image acquisition rate (2 Hz), the false positive rate will drop to 5.31 ± 1.50%.

The neurons recorded *in vivo* in the pretectal area of larval zebrafish had calcium events in 28.5 ± 14.2% of the 12,000 CMN stimulus frames (mean ± standard deviation). On average each neuron generated 158 ± 87.5 event bouts during the 800-second CMN stimulation, *i.e.* each frame-by-frame calcium event was usually part of a longer-lasting calcium event bout. The ROIs in the *in vivo* calcium imaging recording with low averaged event rate (responding in less than 8% of the stimulus frames) were excluded from later analysis (242/1964 ≈ 12.5% ROIs excluded). The binary calcium event trains (2 Hz) of the remaining neurons were up-sampled with the nearest-neighbour resampling method and aligned with the stimulus frames displayed at 15 Hz for reverse correlation.

2.8. Direction preference of motion-sensitive pretectal neurons

To find the ROIs responding to large-field motion stimuli and to get an independent assessment of direction selectivity of these ROIs, the full-field motion stimulus consisting of binary noise patterns moving in 8 directions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) was displayed before the CMN stimulus to the zebrafish larvae in the *in vivo* calcium imaging recordings (Fig. 7a). The stimulus consisted of 3 repeats of 8 phases. Each phase consisted of a 5-second moving period and a 5-second still period. The 8 moving directions were presented in a random order in the moving periods of the 8 phases in each repeat. The spatial frequency of the binary noise pattern in the full-field motion stimulus and the CMN stimulus was the same (0.1 °/cycle).

The linear correlations between calcium fluorescence trace in each phase and the motion state of each phase (moving/not moving) were computed first. The cells that showed strong positive correlation (> 0.7) in more than one of these phases were classified as motion-sensitive pretectal neurons (343 out of the remaining 1651 ROIs). To obtain a detailed direction preference profile of these neurons, the averaged calcium fluorescence intensity in the phases with linear correlation > 0.3 were used as the “direction preference scores”. Since the ratio but not absolute value of these scores was important for determining the directional preference, we normalized these scores to the

range of [0, 1] for each of the 3 repeats. The score for each direction can also be written in a vector form whose phase angle and norm were the moving direction and the normalized direction preference scores respectively. The vectors for average normalized scores in the 3 repeats were summed to calculate the most preferred direction which equalled to the phase angle of the vector sum.

2.9. Reverse correlation

The event-triggered average (ETA) for receptive field estimation was computed as:

$$ETA = \frac{1}{n_{event}} \sum_{i=1}^{I} e_i \boldsymbol{B}(\boldsymbol{V_i})$$

$$n_{event} = \sum_{i}^{I} e_i$$

The e_i indicated the i^{th} state (0 or 1, each state corresponds to a period of 67 ms) in the binary calcium event trains, and the $\mathbf{V_i}$ was the motion vector field at the same moment. The direct summation of motion vectors can be ambiguous because the length of the summed vector is dependent on both the circular variance and the length of the contributing vectors. We therefore used a binned representation, $\mathbf{B(V_i)}$, in which each motion vector v in $\mathbf{V_i}$ was binned as follows into direction bins:

$$\boldsymbol{B}(v) = (a_0, a_1, \ldots, a_{n-1})$$

$$where \quad v = (r, \angle\theta), \ \angle\theta \in [-\pi, \pi]$$

$$and \quad a_i = \begin{cases} r & i = \lfloor \frac{n(\angle\theta + \pi)}{2\pi} \rfloor \\ 0 & otherwise \end{cases}$$

This binning process leads to a sparse, one-hot representation. The phase angles $\angle\theta$ of motion vectors v are evenly divided into n = 16 direction bins; the norm r of v is assigned to the corresponding direction bin. This results in a transformation of the motion vector matrix of the stimulus into an $8 \times 28 \times 12{,}000 \times 16$ matrix (elevation × azimuth × frame × direction bins). The ETA computed from this transformation of the motion vector matrix resulted in an $8 \times 28 \times 16$ binned matrix.

For the computation of the whitened ETA, we proceeded as described above except we used the pre-whitened binned stimulus matrix ($\mathbf{B_{whiten}(V)}$) for this computation. The pre-whitening transformation was performed in a similar manner as described in previous studies (Theunissen et al., 2001; Aljadeff et al., 2013):

$$\mathbf{B_{whiten}}(V) = C^{-\frac{1}{2}}(V)\mathbf{Sp}(V)$$

$C^{-\frac{1}{2}}(V)$ was the inversed square root of the covariance matrix of $\mathbf{B}$(V).

2.10. The two-step nonparametric cluster-based bootstrapping test

The two-step nonparametric cluster-based bootstrapping test (2-step NCB test) in this study was developed based on the nonparametric cluster-based permutation test used in EEG and fMRI data analysis (Maris and Oostenveld, 2007; Nichols and Holmes, 2002). This statistic test aimed to find the ETA components whose statistics depend significantly on the calcium event states. These components were considered RF-related structures by the definition of receptive field (Hubel and Wiesel, 1962). The goal of this 2-step NCB test was to identify the significant ETA components by finding the components which would have been highly unlikely to occur, if visual stimulus and neural activity were unrelated.

The procedure of the 2-step NCB test was as follows:

1 To preserve the temporal structure of the binary calcium event trace, instead of random permutation, the bootstrapping distribution was generated by circular shifting of the event trace with a random distance, D in time (Fig. 5A). Since the temporal correlation of the stimulus may violate the exchangeability assumption of the NCB test, we minimized the correlation effect by setting a minimum shift distance equal to 2 times of the temporal CR (abs(D) > 30 frames = 2 s) of the CMN stimulus. The bootstrapped event trains had the same firing rate and noise statistics as the original event train, but their firing was independent of the stimulus pattern. For each simulated or recorded pretectal neuron, we computed 1000 bootstrapped event trains. For each original and each bootstrapped event train the corresponding ETAs were calculated.

2 The distribution of vector norms of the bootstrapped ETAs was computed for each of the 16 direction bins and each spatial position (8×28). With these empirical probability distributions, the unlikelihood that the ETA vector norms were sampled from the empirical probability distribution (or the empirical p-values), can be estimated as follows for all direction bins and spatial positions in the original and bootstrapped ETAs (also illustrated in Fig. 5B):

$$p = 1 - P(\text{ETA value} \in \text{the bootstrapping distribution}).$$

3 Thresholds (= 0.95) were set at both ends of the distribution of empirical p-values obtained in the 3^{rd} step. To avoid confusion with another threshold in a following step, this threshold is referred to as Bernoulli threshold. The units with supra-threshold p-value were labelled as units of interest for both the original and the bootstrapped neurons. Note the Bernoulli threshold here controls only the sensitivity of the statistic test but not the false positive rate in the multiple comparison test, which is only controlled in the following steps (see below). Hence the supra-threshold unit here should not be considered as a significant unit. The supra-threshold units in the two ends of the empirical p-value distribution were called high-value and low-value units (Fig. 5B, shown as the cyan and red circles). This identification of individual spatial supra-threshold units established the first step of the 2-step NCB test.

4 The supra-threshold ETA units that were spatially connected and/or in adjacent direction bins were grouped into clusters. For each cluster in the original or bootstrapped ETA, we summed the supra-threshold p-values (from step 1 above) of all its units as the cluster-level statistic (Fig. 5C).

5 We selected the maximum cluster-level statistic ("cluster score") for each bootstrapped ETA. The distribution of the selected maximum statistic was called the bootstrapped maximum statistic distribution. Using this distribution, we calculated the empirical cluster p-values ($P_{cluster}$) for all clusters in the original ETAs as follows:

$$P_{cluster} = P(\text{ETA cluster statistic} \in \text{the bootstrapped maximum statistic distribution})$$

The clusters whose $P_{cluster}$ was more extreme than a significance threshold were labelled as significant ETA patches (Fig. 5C).

We repeated step 3–5 for the low-value direction bins. The high-value and low-value significant ETA patches can be regarded as visual field locations and motion directions for which high and low stimulus speeds are associated with calcium events, respectively. The significance threshold in step 5 that controlled the false positive rate of the test was set to 0.025 (*i.e.* a two-tailed test at alpha = 0.05).

2.11. Excitatory-Suppressive index

There were receptive field structures, which down-modulated the firing rate when their preferred stimulus was presented. We refer to

such structures as "suppressive receptive field components" here; *vice versa*, the receptive field structures which up-modulated the firing rate are referred to as "excitatory receptive field components". The majority of motion-sensitive pretectal neurons in larval zebrafish prefer motion speeds > 15 °/second (for a spatial period of 30 °/cycle) (Kubo et al., 2014). We therefore expect that the modulation strength in most of the motion-sensitive neurons should increase with movement speed within the low-speed region of 0–6.7 degree/second (spatial period: 10 °/cycle) of our CMN stimulus. We therefore interpret the positive and negative response modulation slopes in this stimulus speed regime as being related to excitatory and suppressive effects, respectively. For example, a neuronal RF component that has higher activity at a speed of 1 °/second than at a speed of 5 ° per second is interpreted as having a suppressive effect, rather than interpreting it to simply be positively tuned to extremely low speeds.

Since each frame of the CMN stimulus was aligned with the corresponding event state, it was easy to calculate the relationship between the movement speed presented in each stimulus unit (8 × 28 spatial unit positions, 16 direction bins) and the event rate (Fig. 8A, right side, Figure S6). This can be done by sorting the event train as follows: For each spatial and directional unit, the time series of speeds (vector norms) was sorted from low to high. The event rate for each unit can be approximated by the moving average of the speed-sorted event train with a moving window size of 10 stimulus frames (0.67 s). An excitatory-suppressive index (ESI) was calculated as the correlation coefficient between the sorted speed values and the speed-sorted event rate. The units with higher ESI in the significant ETA patches were related to an excitatory receptive field structure, while lower ESI values

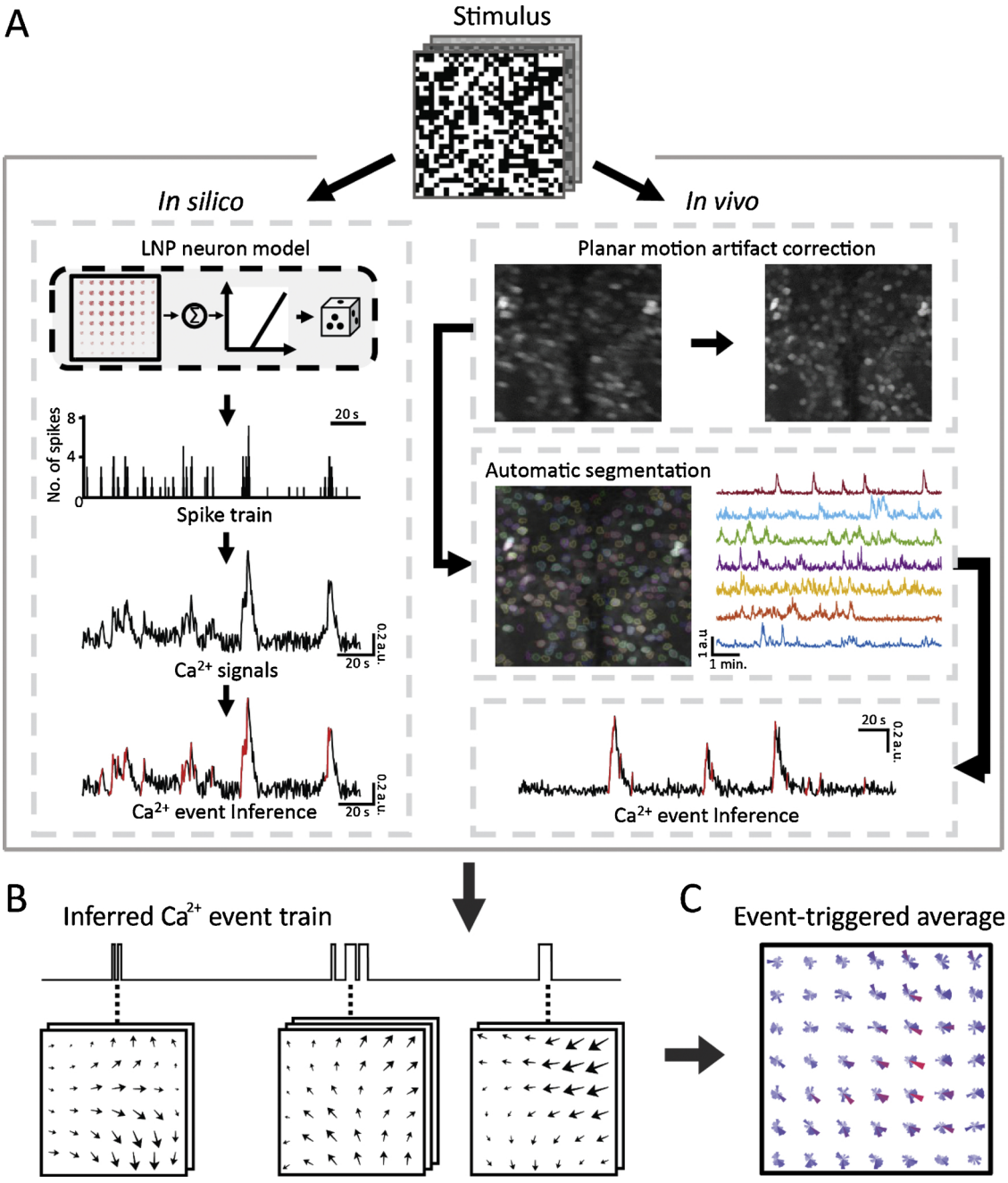


Fig. 3. Analysis workflow for the calculation of calcium event-triggered stimulus averages. **(A)** Preprocessing procedures for the *in silico* (top left) and *in vivo* (top right) experiments, resulting in a time series of calcium events for each neuron; **(B, C)** the event-triggered average is calculated using reverse correlation, resulting in a stimulus motion direction histogram (16 direction bins) at each spatial location; s = second; a.u. = arbitrary units.

corresponded to a suppressive receptive field component.

To visualize directional preference and ESI index, we colour-coded the polar histograms in Fig. 8 and Figure S4 according to the ESI index, while histogram bar length and quiver vector length corresponded to the locally estimated ETA speed as follows: for high-value (red) ETA units, the bar/vector length is proportional to the estimated ETA speed; for low-value (blue) ETA units, the length is inversely proportional to the estimated ETA speed (it corresponds to the difference between two times the averaged stimulus speed and the estimated ETA speed). Thus in Fig. 8 and Figure S4 the histogram scaling is linear, while in Figs. 4 and 6 a square function was used to increase visibility (see explanations in the Figure legends).

3. Results

Efficient characterization of receptive fields (RFs) of visual motion-sensitive neurons is essential for understanding motion processing mechanisms of the brain. The estimation efficiency of motion RFs is mainly limited by the neural activity recording method and the employed stimulus protocols. To account for the poor sensitivity and temporal resolution of the – otherwise promising – calcium imaging method, we developed a novel contiguous motion noise (CMN) stimulus (Fig. 1, Supplementary Video 1) and a bootstrapping-based statistic test. In combination with the motion reverse correlation techniques introduced in previous studies (Borghuis et al., 2003; Srinivasan et al., 1993), the receptive fields of a high number of large-field motion-sensitive neurons can be efficiently estimated in parallel using *in vivo* calcium imaging.

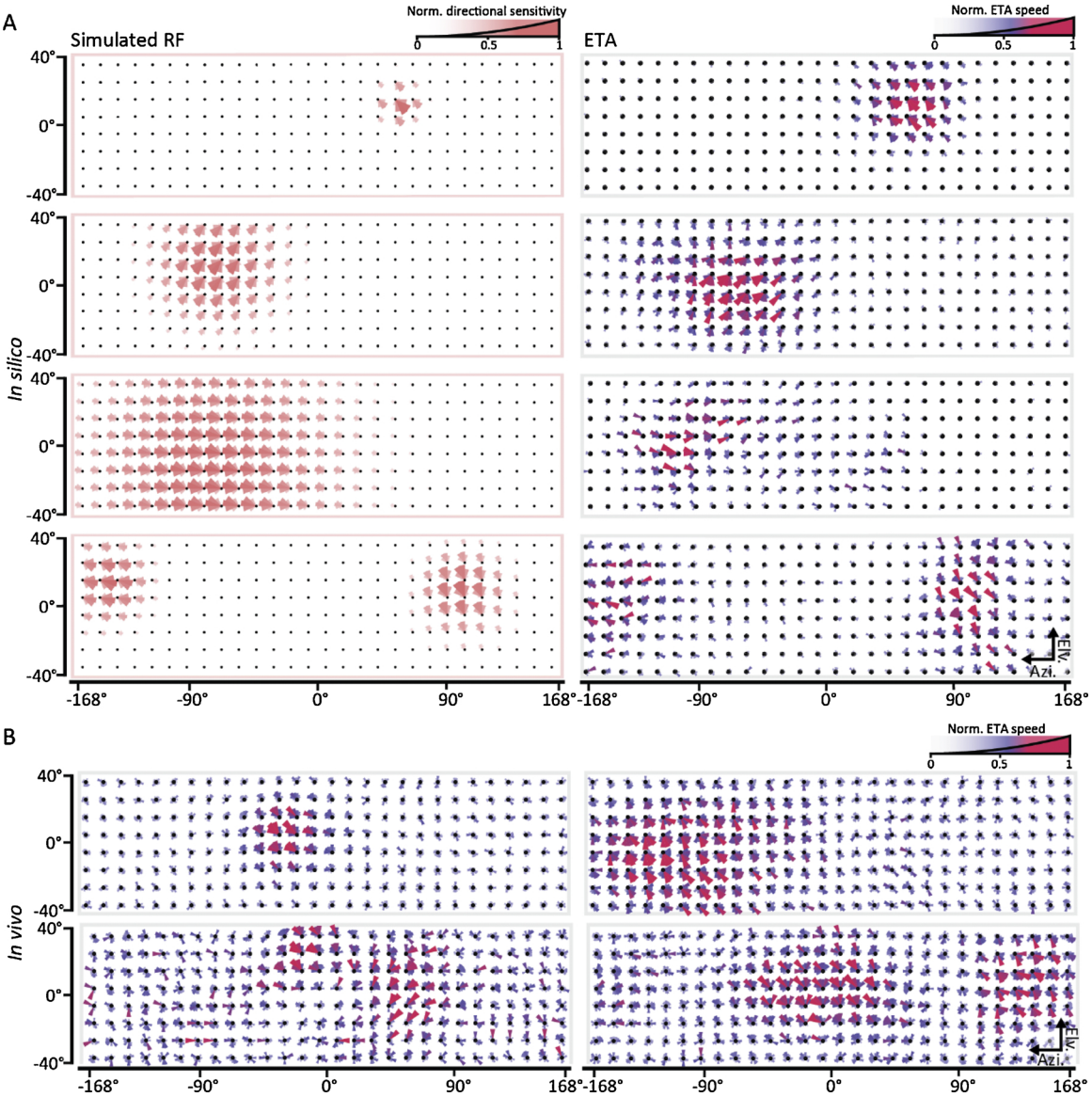


Fig. 4. Estimation of the receptive field structure for simulated (A) and real (B) neuronal calcium responses. **(A)** The ground truth receptive field structure (left column) of four simulated example neurons and the corresponding event-triggered averages (ETAs, right column). From top to bottom, each receptive field patch in the example neurons covered 27°, 82°, 164° and 54°. The bar length in the polar histograms of the ground truth receptive field plot (left column) indicates the local direction preference (normalized to the range of 0–1 by min-max rescaling, 1 indicating the most preferred direction across spatial locations). The colour and the length in the ETA polar histograms (right column) corresponds to the event-triggered averaged speed normalized to the range of occurring speeds in the ETA. In order to increase the visibility of the relevant RF features, the bar lengths were gamma-adjusted using a square function, see non-linear colour legend. Azi. = azimuth; Elv. = elevation. **(B)** Examples of ETAs of four pretectal neurons in a larval zebrafish. Again, the colour and length of individual polar plot histogram bars indicates the squared and normalized ETA speed.

The CMN stimulus consists of spatially and temporally coherent visual field patches (Fig. 1A). In each patch, a binary noise image is presented that moves in a particular direction over time (Fig. 1B), resulting in multiple, spatially tiled motion patches of a certain size, in which binary images translate in different directions. The movement trajectories are generated by first picking random motion directions for each patch and then smoothing them in space and time using a multivariate Gaussian kernel (see Material and methods), which results in coherent motion patterns with smooth motion transitions in-between different stimulus arena positions.

3.1. Statistical properties of the contiguous motion noise (CMN) stimulus

The receptive field structure estimated with the reverse correlation method can be biased if the stimulus contains biases (Chichilnisky, 2001). We therefore first determined whether, and if so, to what extent the relevant feature statistics in our CMN stimulus were biased. First, we characterized the motion direction statistics. If there were biases in the local and global motion direction statistics of the CMN stimulus, the estimated preferred direction in the RF would be misleading. As shown in Fig. 2A, both the local and global statistical distributions of motion direction were mostly circularly uniform over the 12,000 frames of the particular CMN stimulus used in this study. The local distributions at different spatial locations were similar to each other and suggest a spatial uniformity of the direction information in the stimulus.

While the speed information in the stimulus is not critical for interpreting the estimated motion RFs, a heterogeneous stimulus speed distribution may affect neural responses to stimuli displayed in different regions of the visual field. These effects may cause the estimation to have a biased location and/or shape. To evaluate this, the local speed distributions of the CMN stimulus were compared between all spatial locations (Fig. 2B, left). Again, no noticeable difference between these distributions were observed. In addition, the density distribution of speed of all motion vectors contained in the stimulus were fitted well with a modified Rayleigh distribution. According to the fitted function, 95% of the motion vectors in our CMN stimulus had speeds in the range of 0–30 visual degrees/second.

We constrained the spatiotemporal motion correlation in our CMN stimulus to present a sufficient level of contiguous motion, which is regularly present in natural scenes and should lead to strong activation of global motion sensitive neurons. We defined the contiguous radius (CR, see Materials and methods), which corresponded to the radius in which motion correlation was smaller than 0.1 in the cross-correlograms. As shown in Fig. 2C, the spatiotemporal correlation in our CMN stimulus was defined by the spatial and temporal CRs (57 visual degrees and 1 s, respectively). The CRs need to match the RF structures of interest, since choosing very large CRs will prohibit detection of fine-scale spatiotemporal RF structure and very small CRs will not trigger calcium

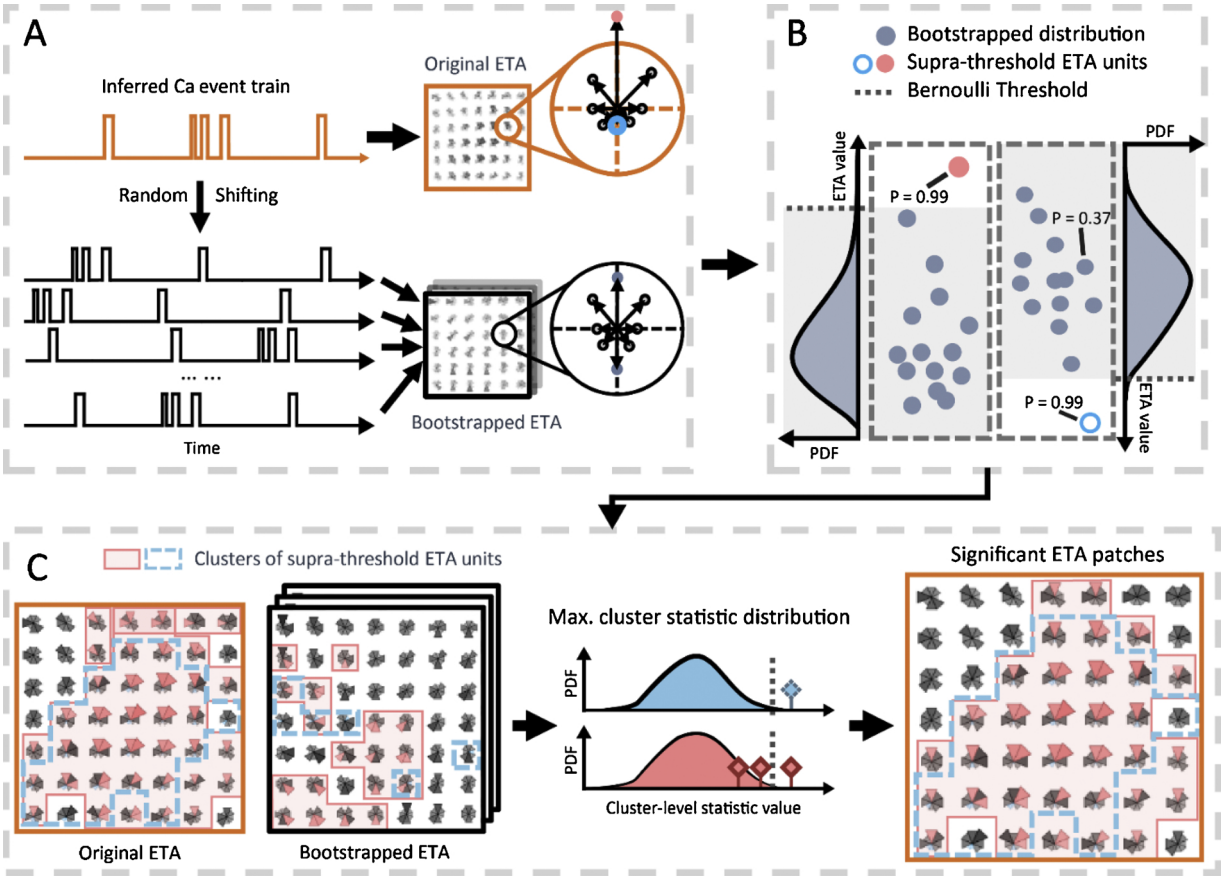


Fig. 5. Workflow of the two-step non-parametric cluster bootstrapping test (two-step NCB test) used to identify the significant ETA patches. **(A)** The bootstrapped ETAs were computed from the randomly circularly shifted calcium event trains in time for a given neuron in question. **(B)** Based on the bootstrapping distribution for each motion direction bin, the empirical probability values (p-values) were computed for each measured and bootstrapped ETA unit (a direction bin of a local ETA polar histogram). The dashed lines indicate the thresholds set for selecting supra-threshold units in the bootstrapping distribution for the two directions. **(C)** Supra-threshold ETA units from the same end of the bootstrapping distribution in B that were spatially connected or in adjacent direction bins, were assigned to the same cluster. The cluster-level statistic was calculated as the sum of P values of all units in each cluster. The distribution of the maximum cluster-level statistic for each bootstrapped ETA was computed as the cluster-level null distribution. The ETA clusters with more extreme values than the significance threshold set in the null distribution (dashed line in the middle plot) were classified as significant ETA patches. A more detailed explanation of the 2-step NCB test is provided in the Methods section.

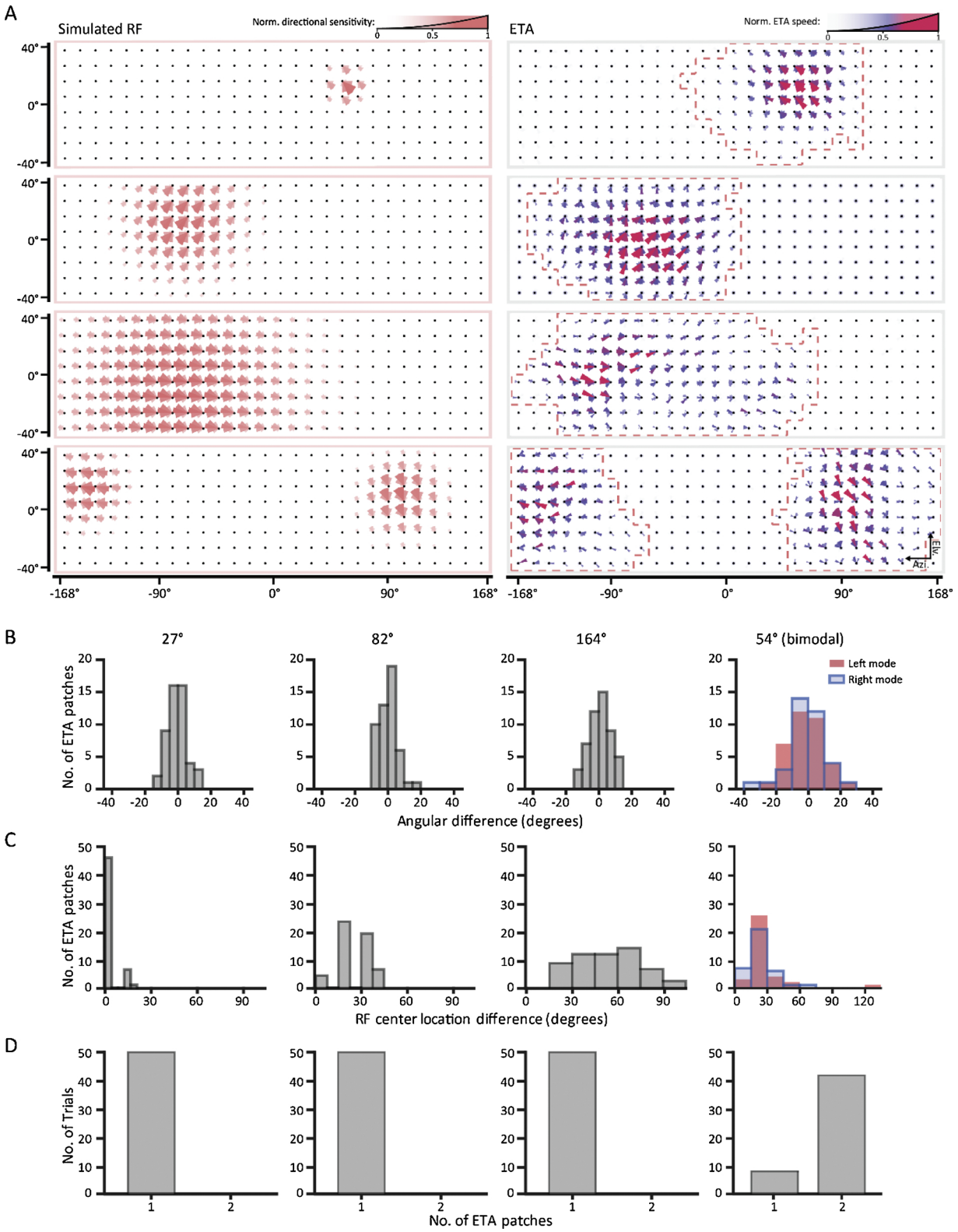
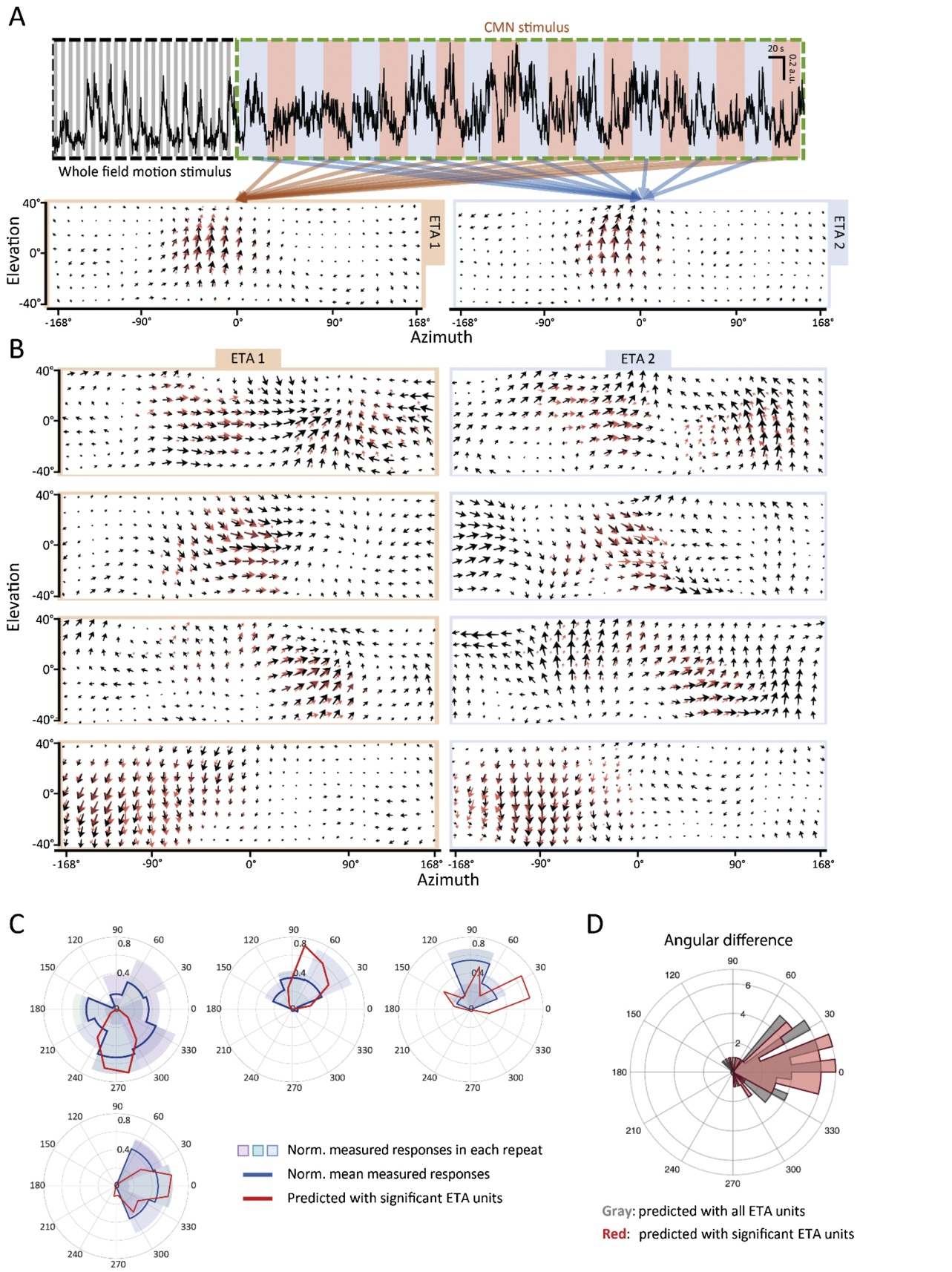


Fig. 6. Performance of our receptive field estimation method in the simulated motion-sensitive neurons. **(A)** The ground truth receptive field structures (left column) of the four simulated neurons shown in Fig. 4A and the corresponding significant ETA patches identified as the estimated receptive field structures (outlined with the contours in dashed red in the right column plots). The bar length in the polar histograms of the ground truth RFs and ETA plots corresponds to the square of the normalized direction preference and event-triggered averaged speed, respectively (as in Fig. 4). Azi. = azimuth; Elv. = elevation. **(B)** The angular difference between the ground truth preferred direction of the simulated neurons and the preferred direction predicted by the significant ETA patches for simulated neurons with different monocular receptive field sizes (27°, 82°, 164°) or binocular receptive fields containing one receptive field patch (54°) in each the left and right halves of the visual field (n = 50 neurons for each type of receptive field). For the binocular neurons, the histogram for the preferred directions of receptive field components in the left and right visual fields were labelled in red and blue respectively. **(C)** The difference between the ground truth receptive field centre locations and the centre locations of the significant ETA patches. **(D)** The number of spatially isolated significant ETA patches identified in the simulated monocular and binocular neurons.



(caption on next page)

Fig. 7. Receptive field estimation performance in the pretectal neurons recorded by *in vivo* calcium imaging. **(A)** An illustration of the *in vivo* calcium imaging stimulus protocol and analysis. The neuronal direction selectivity was measured by the responses to whole-field motion stimuli (three repetitions of binary noise patterns moving in 8 different directions, see Materials and methods for details) before the CMN stimulus. The CMN stimulus was divided into two interlaced subsets of equal length, and for each set an ETA was computed independently. The preferred motion direction for each spatial location in the ETAs computed for an example *in vivo* pretectal neuron is indicated by the black arrows in the quiver plots. The red arrows represent the local direction preference of the significant ETA patch computed using the whole CMN stimulus period. **(B)** Same analysis as in A for four example pretectal neurons. **(C)** The averaged peak calcium response amplitude for the whole-field motion stimuli is shown as blue lines in the four example neurons shown in B, in comparison to the peak response amplitude predicted by the estimated receptive field structures (red). The areas shaded in different colours correspond to the response amplitudes in the three repetitions of the whole-field motion stimuli. The measured and predicted peak response amplitude was normalized to the range of 0–1 by the maximum response amplitude of the average direction selectivity measured (see Materials and methods for details). **(D)** The angular difference between the measured preferred direction and the predicted preferred direction based on the ETA (grey) or the significant ETA only (red) in 163 motion-sensitive neurons recorded in a larval zebrafish.

events at sufficient frequency.

The time-averaged luminance for each pixel in the binary movie of the CMN stimulus was also computed (Figure S1). No strong biases were observed in the averaged luminance plot.

3.2. The reverse correlation approach for simulated and recorded neurons

We first estimated the motion RF structure using the reverse correlation technique. As a classic receptive field estimation method, the reverse correlation technique has already been applied to estimate a wide range of RF structures in the past, including motion RFs (Borghuis et al., 2003; Srinivasan et al., 1993). In most of these studies, the neuronal output was recorded as action potential spikes *via* electrophysiology setups, and the RF structures were estimated as the spike-triggered average (STA) pattern. Since calcium activities corresponding to single action potentials are difficult to detect, we developed our RF estimation to depend on the calcium activities triggered by significant increases of neuronal firing rate instead; and the motion RF was thus estimated as the calcium-event-triggered average (ETA) (Fig. 3).

The calcium events in the calcium response traces in pretectal neuronal somata were detected by a Gaussian-mixture-model-based algorithm (Fig. 3, see Materials and methods for details on data preprocessing and calcium event detection). The calcium event train was aligned with the CMN stimulus in time to establish the correspondence between calcium event state (with/without event detected) and current stimulus frame. The ETA corresponded to the average of those CMN motion vector stimulus frames that coincided with calcium events. To maximally preserve the information about stimulus speed and direction, we used a binning transformation of motion vectors for our ETA, *i.e.* at every stimulus location we calculated an ETA histogram of stimulus speeds for each of 16 equally spaced directional bins (Fig. 3C).

We wanted to test the performance of the reverse correlation method. Naturally, the ground truth motion RF structures of neurons recorded *in vivo* are not available. Our estimation method was therefore tested with simulated neurons with known ground truth receptive fields in an *in silico* experiment, where different sources of noise and slow calcium signal kinetics of the *in vivo* calcium recordings were mimicked (Fig. 3A, left). The simulated neurons were constructed as linear-non-linear-Poisson (LNP) models (see Materials and methods), whose linear filters were considered as the ground-truth motion RF. These designed receptive fields were of Gaussian shape (different diameters) and each simulated neuron contained either one or two Gaussian patches. The same calcium event detection algorithm (see Materials and methods) as for the *in vivo* data was then used to detect the events in the simulated calcium responses, and thus the event detection error caused by different noise sources and experimental limitations was included in these simulations.

As shown in Fig. 4 A, the locations and the preferred directions of the ground-truth RFs of the simulated neurons were well estimated by the ETAs. Furthermore, for our *in vivo* neurons (Fig. 4B), ETA structures could be identified that differed across neurons and showed different levels of organization within individual neuronal ETAs, suggesting that the ETAs were related to the receptive fields of direction-selective pretectal neurons (Wang et al., 2019).

3.3. Identification of significant event-triggered average patches

The genuine RF-related structures within the ETA need to be identified for any further quantitative characterization of the receptive field properties. The identification of RF-related ETA structures will be less biased if no or very few assumptions are made in the identification method. Here, we only assumed the motion receptive field structures to be locally contiguous. With this assumption, we employed a two-step nonparametric cluster-based bootstrapping test (or 2-step NCB test, Fig. 5, see Materials and methods for details) to identify the ETA units clustered in space and/or in adjacent directional bins, which were significantly related with neuronal responses. In the first step, calcium event trains were circularly shifted in time to obtain bootstrapped ETAs, in which the shifted calcium events were no longer related to the CMN stimulus frames. This step identified motion directions across visual field locations for which the ETA stimulus speeds deviated from chance speeds (Fig. 5A-B). In the second step, the algorithm identified spatial-directional clusters of ETA units making up one significant ETA patch (Fig. 5C). The significant ETA patches were identified by comparing the cluster score (a measure of the spatial-directional extent of the patch, see Materials and methods) of an ETA of an *in vivo* or simulated neuron to the maximal cluster scores of the bootstrapped ETA clusters for that neuron. Each identified significant cluster either contained units with unusually high stimulus speeds, or units with unusually low speeds, so that clusters with antagonistic properties were kept separate (red and blue colours in Fig. 5C).

The performance of our 2-step NCB test was quantified in *in silico* experiments, where the estimated significant ETA patches of simulated neurons were compared to their ground truth receptive field structures. As shown in Fig. 6 A, the RF- related structures in the ETAs of the simulated neurons shown in Fig. 4A were correctly identified by the 2-step NCB test as the significant ETA patches, which are plotted as the spatially connected direction bins outlined by a red dashed line. To quantitatively evaluate the estimation accuracy, the estimation error for the RF centre locations and the preferred directions of the RF structures were computed for the simulated monocular neurons with one Gaussian component, whose radius varied from 27 to 164 visual degrees (n = 50 simulated neurons for each radius size). The estimation error of the preferred direction was less than 20 degrees for all simulated neurons (Fig. 6B). The RF centre location was reliably detected for small-size RF neurons, but oftentimes offset by 30° or more for large-size RF neurons (164° radius) (Fig. 6C). The averaged RF centre location errors for simulated RF radii of 27°, 82° and 164 visual degrees were 2.25°, 24.3°, and 50.7°, respectively, which correspond to 7.5%, 27% and 28% of the respective RF radius. This showed that the centre position for the large simulated RFs were in general relatively difficult to localize. Since our LED arena only covered part of the visual field, and this limitation was faithfully included in the *in silico* experiments, the receptive field outside of the arena-covered region cannot be estimated. These non-estimated parts may cause the localization problem, especially when the receptive field is large, which corresponds to a higher RF proportion being outside of the covered region. We also computed the estimation errors for preferred direction and RF centres for simulated binocular neurons that had two Gaussian RF components. The

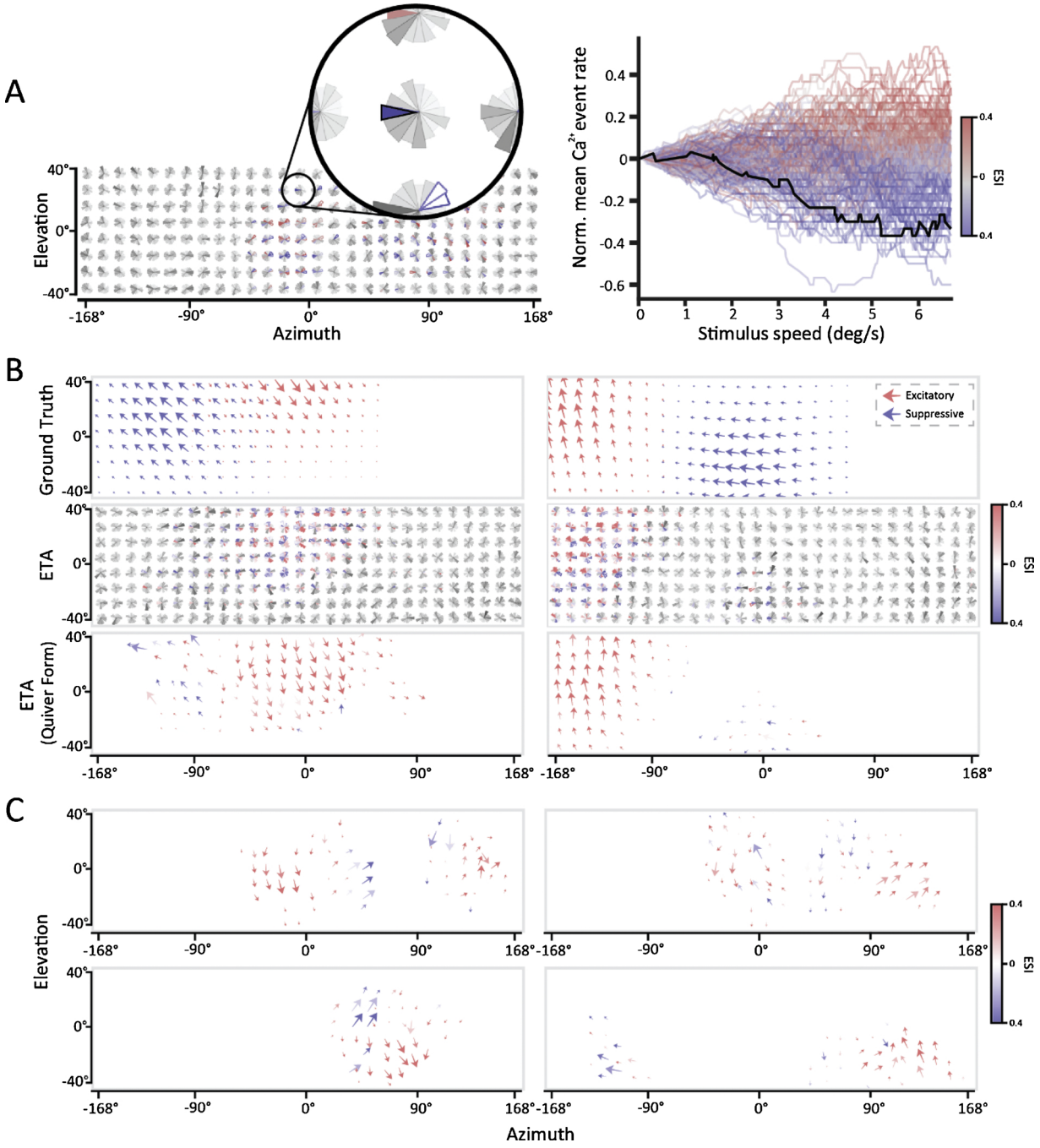


Fig. 8. Estimation of suppressive receptive fields using the CNM stimulus and the directional motion speed tuning of neurons. **(A)** Assignment of an excitatory-suppressive index (ESI) for each ETA unit based on the correlation of speed and calcium event rate (see Materials and methods and Supplementary Figure S6 for details). To make low-value ETA units visible in the polar plots, they were outlined with bars whose length equalled two times the mean CMN vector norms of the corresponding positions minus the vector norms of the directional units (which increased the length of the low-value unit bars relative to other bars). Each curve in the right plot corresponds to a unit in the significant ETA patches (direction bin) in the left plot. For example, the black curve in the right plot corresponded to the direction bin labelled with thick black border in the left plot. The curve colour corresponds to the ESI value. **(B)** The comparison between the ground truth receptive field structure of two simulated neurons (top) and the corresponding significant ETA patches (middle: shown in polar histogram form; bottom: shown in simplified quiver form). The red and blue arrows in the top subplots represent the excitatory and suppressive receptive field, respectively. The significant ETA patches were coloured by the corresponding excitatory-suppressive index (ESI). Non-signifcant ETA patches are coloured in grey. **(C)** Estimated receptive fields of four pretectal neurons recorded *in vivo* are shown in quiver form.

estimation error of the preferred directions for the majority of these simulated neurons was less than 20 degrees (Fig. 6B) and the normalized average centre location estimation errors were 34% of the receptive field radius (Fig. 6C).

To evaluate the false negative and false positive detection of the receptive field components, we counted the number of detected significant ETA patches per neuron: For all simulated monocular neurons exactly one significant ETA patch was detected, corresponding to the single Gaussian RF component which these neurons were designed to have (Fig. 6D). For the simulated binocular neurons, the two Gaussian components were successfully detected in 84% of the simulated neurons (Fig. 6D). No false positive detection was encountered for any

of the simulated neurons.

In addition, we ran simulations to compare our method with three other representative RF mapping techniques used in previous studies, including: the filtered back-projection technique (Fiorani et al., 2014; Johnston et al., 2014), the motion reverse correlation technique using vector white noise stimulus (Borghuis et al., 2003; Srinivasan et al., 1993) and the method developed by Krapp and Hengstenberg (1997), which estimated elaborate motion receptive field structures using the neural responses to circularly moving dots across different spatial locations. The signal-to-noise ratio, kinetics and sampling rate of the simulated calcium signals were the same for all tested RF estimation methods (see Materials and methods section). The stimulus protocol duration was the same across methods (except for the local motion method, which needed at least 320 min for a complete estimation). Due to these design adaptations, some details of the implemented methods (*e.g.* data preprocessing, the chosen representative parameters) may differ from the original studies.

As shown in Figure S2, our method ("CMN + 2step NCB method") successfully estimated the receptive field structures of all simulated neurons, while the other methods failed to estimate the large-size or the binocular RFs. Thus, in our simulation of different calcium imaging RF estimation techniques, our CMN method outperformed the other implemented methods.

Considering that the signal-to-noise ratio of *in-vivo* calcium signals can vary across neurons and experimental conditions, we evaluated the performance of our approach under 3 different noise levels (Figure S3). Although the ETA estimation quality dropped at increased noise levels, the 2-step NCB statistic test still correctly identified major portions of the ground truth RF. This suggests that the 2-step NCB test contributes significantly to the noise-robustness of our approach.

3.4. Motion receptive field estimation in vivo

In order to put our CMN approach to test *in vivo*, we recorded GCaMP6s calcium signals in the pretectum of larval zebrafish. In 163 out of 343 identified motion-sensitive pretectal neurons, RF-related significant ETA patches were successfully identified by the two-step NCB test. It is possible that RF-related structures also exist in the ETAs of the remaining neurons, however there was not enough evidence for the statistic test to discriminate these structures from the noise and biases. To evaluate the performance of the motion RF estimation and get an independent measure of directional preference for each neuron, we included whole-field motion stimuli of different directions in the beginning of the recording (Fig. 7 A). For each *in vivo* neuron, the directional preference was predicted by using all ETA units in its significant ETA patch. The predicted directional preference was then compared to the preferred direction measured using the full-field motion stimuli (Fig. 7A, see Materials and methods for details). For the majority of neurons, the predicted and measured directional preference matched well, resulting in angular errors smaller than ± 30 ° (Fig. 7C & D), and showing no bias in the prediction.

As an additional test for the reliability of our method, we divided the CMN stimulus into two interlaced, half-length stimulus sets with equal number of frames, and compared the two ETAs estimated for the same neuron using these two stimulus sets. We reasoned that if our method resulted in valid RF estimations, the two ETAs estimated with these stimulus sets for a given neuron should strongly resemble each other. As expected, these invariant ETA structures were found for almost all neurons; moreover, for a given neuron, these invariant structures were largely overlapping with the significant ETA patches identified using the full-length CMN stimulus (Fig. 7A & B). Note that the examples in Fig. 7A & B are shown in quiver plot form instead of the polar-histogram form used in previous figures only for illustrative purposes (both types of plots are included for eight additional neurons in Figure S4). Together with the analysis of directional preference, our observations suggest that reverse correlation with the CMN stimulus in combination with the two-step NCB test can be used to estimate the motion RF structure of pretectal neurons efficiently and accurately from *in vivo* calcium recordings. In addition, we also computed the ETAs with a pre-whitened CMN stimulus (Figure S5). The structures in the whitened ETAs were mostly consistent with the un-whitened ETA structure but noisier.

3.5. Estimation of suppressive receptive fields

The significant ETA patches in the *in silico* and *in vivo* estimations presented in Figs. 6 and 7 corresponded to those ETA clusters for which the 2-step NCB test had detected unusually high motion speeds, *i.e.* the high-value clusters. The motion units related to low-value clusters, however, had low or no speed. It seemed likely that these low-value ETA clusters were related to the suppressive components in the receptive field. To infer the suppressive RF components, we therefore modified and used the categorical reverse correlation method introduced in a previous study (Ringach et al., 2003), which makes use of both the high-value and low-value clusters. The principle of this method was to compute the correlation between event rate (firing rate) and the stimulus intensity (speed) for each unit within the ETA. A more positive correlation suggested the ETA units were more related with the excitatory RF structures, and a more negative correlation suggested relation to suppressive RF structures (see Materials and methods for details, Figure S6). Based on this principle, we computed an excitatory-suppressive index for all units in the significant ETA patches containing both the high- and low-value clusters (Fig. 8A). Again, this inference method was first tested in the *in silico* simulation. The suppressive RF components were successfully inferred by this method in 21 out of 50 simulated neurons tested. Examples of these suppressive components are shown in Fig. 8B. Note that these simulated neurons did not have a background firing rate, which meant their firing rate would be 0 when the excitatory RF components were not triggered. Suppressive components are thus more difficult to detect than excitatory ones when only positive activity levels can be measured – as is the case for calcium imaging.

In Fig. 8C, example *in vivo* neurons with estimated excitatory and suppressive RF structures are shown (colour-coded by the excitatory-inhibitory index, ESI, see Materials and methods). We found the motion RF structures of many pretectal neurons to be more complicated than the Gaussian structure we designed for the simulated neurons (more examples in Figure S4): the shapes of these receptive field structures were relatively irregular; and the direction preference profile in different local regions in the receptive field showed big variations.

In addition, we also estimated the spatiotemporal RF structure for luminance features for these pretectal neurons by computing the averaged pattern of the event-triggered binary noise image frames of the CMN movie. However, no spatiotemporal structure that could be related to the motion RF structure or the direction preference was observed in these neurons (Figure S7).

4. Discussion

In the visual systems of a wide range of animals, from flies, fishes to mammals, local motion features are integrated into global motion features at different stages in the visual pathways, and hundreds to thousands of neurons are involved in this integration process (Chen et al., 2008; Krapp and Hengstenberg, 1996; Kubo et al., 2014; Kühn and Gollisch, 2019; Portugues et al., 2014; Xiao and Frost, 2013). Estimating receptive fields with high efficiency is essential for the systematic functional characterization of a large number of neurons, especially the complex neurons integrating motion over large parts of the visual field. In this study, we developed a motion receptive field estimation method suitable for calcium imaging that combines a novel motion noise stimulus (CMN stimulus) with the reverse correlation technique, and assesses significant RF structures using a statistical test

(2-step NCB test). With this method, we can estimate the fine-scale receptive field structure for dozens of zebrafish pretectal neurons in parallel in less than 14 min.

The efficiency of motion RF estimation methods is mainly limited by the acquisition rate of RF information and the throughput of the neural activity recording method. The RF information acquisition rate depends not only on the stimulus used, but also on the quality of the neural activity signal recorded. In comparison to electrophysiological recordings, calcium imaging has a high throughput but a limited per-neuron information rate due to the poor sensitivity and temporal resolution of this method. To compensate for this limitation, the stimulus used for receptive field estimation must be optimized to efficiently acquire relevant receptive field information from neuronal responses.

The ideal stimulus should be able to maximize the mutual information between the stimulus and the neuronal responses in a certain period of time (Borst and Theunissen, 1999). This mutual information rate is low when the stimulus is too simple (not informative) or too complex (cannot trigger neuronal responses efficiently). A common strategy for finding the optimal stimulus is to constrain a white noise stimulus with assumptions based on the prior knowledge about the functional properties of the neurons of interest. For example, by assuming a homogeneous direction preference in each part of the receptive field, the RF locations can be quickly mapped by back-projecting moving bar stimuli weighted by neuronal responses to the corresponding location in the visual field (Fiorani et al., 2014; Johnston et al., 2014). However, the assumptions made may also limit the type of receptive field the method can estimate. Our receptive field estimation method was developed based on the motion reverse correlation method developed by Srinivasan and colleagues (1993) and elaborated by Borghuis and colleagues (2003). In our method it is assumed that direction selectivity is locally coherent within the receptive field in space and time. Furthermore, it uses a simple binary noise luminance texture and as such puts less weight on the neuronal tuning to luminance features such as spatial frequency. These simplifications and parameter space reductions provide power for discovering more complex motion RF structures across the visual field. For example, we could demonstrate elaborate bipartite receptive fields with different directional preferences in each of the two RF components (Fig. 8C, S4), many of which sample information from both eyes (*i.e.* they integrate motion binocularly). Moreover, the 2-step NCB test we developed can further exclude the effect of noise sources and stimulus biases on the estimated receptive field structures (Fig. 6, S3). Our method is therefore compatible with *in vivo* two-photon calcium imaging, where activity traces are oftentimes noisy. This enables systematic functional characterizations of large numbers of motion-sensitive neurons, and may therefore facilitate investigation into global motion integration processes at the population level. For example, it could help to determine to what extent the population statistic of the large-size motion receptive fields is similar to the statistics of optic flow fields in natural scenes, which is an important step for understanding the visuomotor transformations underlying stabilization behaviours (Naumann et al., 2016; Wang et al., 2019).

Although the constraint of spatial and temporal frequency in the noise stimulus may help to improve the estimation efficiency to a high extent, the spatiotemporal correlation introduced by this constraint may break the spherical symmetry of the stimulus, which is important for an unbiased estimation with the reverse correlation technique (Chichilnisky, 2001; Paninski, 2003). One bias we noticed is that the high frequency components in the receptive field structures, whose size is smaller than the contiguity radius, may not be recovered in the estimation (*e.g.* the first example shown in Fig. 6A). This bias is likely to have only minor effects on the estimation of the spatial structure of large-size motion RFs, since such high-frequency structures may be less important for global motion integration. Furthermore, the spatiotemporal contiguity of the stimulus limits the estimation of the temporal receptive field structure. Even though we observed temporal structure in the ETAs of many recorded neurons (Figure S8), any possibly existing high frequency temporal RF components are unlikely to be estimated correctly by the CMN method.

One solution to this problem is to repeat the estimation with a stimulus with smaller CR so that small spatial receptive field structure and/or temporal structure can be estimated better. However, such a stimulus may be inefficient in estimating large receptive fields, which may result in longer estimation time or poorer estimation quality. In addition, the spherical symmetry of the correlated stimulus can be restored by a whitening transformation (Schwartz et al., 2006). However, the ETAs computed from the whitened stimulus do not provide more information about the high frequency RF structures (Figure S5), which could be explained by the amplification of noise in the whitening process.

Furthermore, the spatiotemporal correlation of the motion features, the spatial frequency, and the moving speed of the binary noise texture used to display motion could also affect the estimated results. More complex stimuli that involve a richer set of parameters, *e.g.* including a wider range of spatial frequencies and speeds, would be needed for investigating the effect of these features in motion RF estimation. However, the relevant features may be “diluted” in a more complex stimulus, which in turn might reduce the efficiency of triggering strong neuronal responses. As a result of the limited information rate during neural activity recordings, a trilemma is formed by the aspired short estimation time, the aspired high estimation quality, and the implemented estimation of vivid details (spatio-temporal combinations of flow directions). This trilemma does not only exist in our approach but also in the classic RF estimation techniques (Figure S2).

Reverse correlation is limited to estimate the linear structure of receptive fields (Boer and Kuyper, 1968). For the receptive fields with suppressive and excitatory RF components, the reverse correlation method has very limited power to estimate their structures. This limitation results from the suppressive component being nonlinear, since it is rectified to only down-modulate neural activities. These nonlinear structures can be investigated by spike-triggered covariance analysis, which can reveal specific non-linear structures in the receptive field (Schwartz et al., 2002). However, the spike-triggered covariance analysis requires the stimulus to be spherically symmetric which is not satisfied by the CMN stimulus due to its spatiotemporal correlation. Although a previous study proposed a spike-triggered covariance method for a strongly correlated Gaussian stimulus (Aljadeff et al., 2013), it is difficult to apply this method to the estimation of motion receptive fields, which is a vector field but not a scalar matrix.

The method we used to detect the suppressive component from the calcium event-triggered ensemble is modified from the categorical reverse correlation method developed in the previous studies investigating orientation-selective RF structures (Ringach et al., 2003, 2002). In these studies, the spatial frequency tuning properties of the orientation-selective receptive fields interfered with the suppressive component detection. This problem was solved by comparing the firing rate to the baseline firing rate measured by displaying orientations with high spatial frequency textures that these neurons cannot resolve (Ringach et al., 2002). A similar approach is used in our study: the excitatory/suppressive properties of the direction-selective RF structures have been separated from speed tuning properties of motion-sensitive pretectal neurons using low-speed range, which few pretectal neuron are tuned to (Kubo et al., 2014).

The simulated neurons we used to test the inference method for suppressive components had baseline firing rates (or spontaneous activity rate) of 0 Hz. The information of suppressive RF components can only be inferred from the stimulus frames in which a suppressive component is active, while an excitatory component is active as well. Such stimulus frames are of course much less frequent than the frames containing the excitatory component alone. Therefore, our method is less sensitive in inferring suppressive RF components than excitatory components for neurons which are not spontaneously active. Given the

difficulty to infer suppressive components, it is interesting that for the neurons recorded *in vivo*, suppressive components could oftentimes be inferred (Figs. 8C and S4). This could suggest that the *in vivo* neurons had higher baseline activity or stronger suppressive components than the simulated neurons we designed for benchmarking purposes with no baseline activity.

Another issue regarding suppressive component inference is that we cannot confirm the validity of the suppressive components that spatially overlapped with excitatory components (Fig. 8B). The negative correlation between event rate and increasing speed in these components can be caused by the presence of a suppressive receptive field structure, or the absence of stimulus patterns preferred by the excitatory component. Since we don't have any good reason to exclude the second possibility, we decided to exclude the overlapping components from the suppressive component inference.

The spatiotemporal luminance RF structure of many visual neurons clearly corresponds to the local motion they are tuned to as predicted by the spatiotemporal energy model (Adelson and Bergen, 1985; Eckhorn et al., 1993; Wang and Yao, 2011). However, for many global motion sensitive neurons, the correspondence between the luminance RF structure and the motion feature these neurons are tuned to is not clear (Kühn and Gollisch, 2019; Matsumoto and Tachibana, 2017). It was proposed that motion features and static features would be separated to avoid ambiguity in the feature extraction process in the later visual processing (Kühn and Gollisch, 2019). In the CMN stimulus, the motion feature is uncoupled from the static feature. In the estimated motion receptive fields of pretectal neurons, we did not observe any spatiotemporal luminance structure that can be related to the motion RF in any of the pretectal neurons recorded. It is possible that motion features are separated from texture features by the pretectal neurons in zebrafish or already further upstream in the retina. However, an alternative explanation for the observed texture-invariance is that the spatiotemporal luminance structure might have been masked by the noise and low temporal resolution of the recorded signals. Future studies are needed to investigate this question of feature separation in the non-cortical visual pathways of zebrafish.

5. Conclusion

In this study, we established a rapid, robust estimation method for motion receptive fields, which is compatible with high-throughput calcium imaging at low temporal resolution. Using this method, we have successfully estimated complex large-size receptive field structures of motion-sensitive pretectal neurons in larval zebrafish. The estimated receptive fields provide detailed profiles of local direction selectivity in the receptive field (flow fields), and using this local information, we were able to predict the direction preference of these neurons during whole-field motion stimulation. Our results suggest that our CMN method is suitable for the systematic functional characterization of large-size RFs in vertebrate visual brain areas involved in motion processing, such as the zebrafish pretectum.

Funding

This work was funded by the Deutsche Forschungsgemeinschaft (DFG) grants EXC307 (CIN – Werner Reichardt Centre for Integrative Neuroscience) and INST 37/967–1 FUGG, and a Human Frontier Science Program (HFSP) Young Investigator Grant RGY0079.

Declaration of Competing Interest

All authors declare that they have no competing interests.

Acknowledgements

We thank Thomas Nieß (glassblower shop, University of Tübingen) and Klaus Vollmer (fine mechanics workshop, University of Tübingen) for technical support, and Herwig Baier (MPI of Neurobiology) for providing the *Tg(elavl3:nls-GCaMP6s)mpn400* zebrafish line. We thank Fabian Sinz (University of Tübingen) and members of the laboratory for feedback on the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jneumeth.2019.108366.

References

Adelson, E.H., Bergen, J.R., 1985. Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A, JOSAA 2, 284–299. https://doi.org/10.1364/JOSAA.2.000284.

Ahrens, M.B., Orger, M.B., Robson, D.N., Li, J.M., Keller, P.J., 2013. Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nat. Methods 10, 413–420. https://doi.org/10.1038/nmeth.2434.

Akerboom, J., Chen, T.-W., Wardill, T.J., Tian, L., Marvin, J.S., Mutlu, S., Calderon, N.C., Esposti, F., Borghuis, B.G., Sun, X.R., Gordus, A., Orger, M.B., Portugues, R., Engert, F., Macklin, J.J., Filosa, A., Aggarwal, A., Kerr, R.A., Takagi, R., Kracun, S., Shigetomi, E., Khakh, B.S., Baier, H., Lagnado, L., Wang, S.S.-H., Bargmann, C.I., Kimmel, B.E., Jayaraman, V., Svoboda, K., Kim, D.S., Schreiter, E.R., Looger, L.L., 2012. Optimization of a GCaMP calcium indicator for neural activity imaging. J. Neurosci. 32, 13819–13840. https://doi.org/10.1523/JNEUROSCI.2601-12.2012.

Aljadeff, J., Segev, R., Ii, M.J.B., Sharpee, T.O., 2013. Spike triggered covariance in strongly correlated gaussian stimuli. PLoS Comput. Biol. 9, e1003206. https://doi.org/10.1371/journal.pcbi.1003206.

Baden, T., Berens, P., Bethge, M., Euler, T., 2013. Spikes in mammalian bipolar cells support temporal layering of the inner retina. Curr. Biol. 23, 48–52. https://doi.org/10.1016/j.cub.2012.11.006.

Barlow, H.B., Hill, R.M., Levick, W.R., 1964. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. J. Physiology 173, 377–407. https://doi.org/10.1113/jphysiol.1964.sp007463.

Boer, E.D., Kuyper, P., 1968. Triggered correlation. IEEE Trans. Biom. Engin. BME 15, 169–179. https://doi.org/10.1109/TBME.1968.4502561.

Bonin, V., Histed, M.H., Yurgenson, S., Reid, R.C., 2011. Local diversity and fine-scale organization of receptive fields in mouse visual cortex. J. Neurosci. 31, 18506–18521. https://doi.org/10.1523/JNEUROSCI.2974-11.2011.

Borghuis, B.G., Perge, J.A., Vajda, I., van Wezel, R.J.A., van de Grind, W.A., Lankheet, M.J.M., 2003. The motion reverse correlation (MRC) method: a linear systems approach in the motion domain. J. Neurosci. Methods 123, 153–166.

Borst, A., Theunissen, F.E., 1999. Information theory and neural coding. Nat. Neurosci. 2, 947–957. https://doi.org/10.1038/14731.

Chen, A., Gu, Y., Takahashi, K., Angelaki, D.E., DeAngelis, G.C., 2008. Clustering of self-motion selectivity and visual response properties in Macaque Area MSTd. J. Neurophysiol. 100, 2669–2683. https://doi.org/10.1152/jn.90705.2008.

Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., Looger, L.L., Svoboda, K., Kim, D.S., 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300. https://doi.org/10.1038/nature12354.

Chichilnisky, E.J., 2001. A simple white noise analysis of neuronal light responses. Network 12, 199–213. https://doi.org/10.1080/net.12.2.199.213.

Dal Maschio, M., Donovan, J.C., Helmbrecht, T.O., Baier, H., 2017. Linking neurons to network function and behavior by two-photon holographic optogenetics and volumetric imaging. Neuron 94, 774–789 e5.

DeAngelis, G.C., Ohzawa, I., Freeman, R.D., 1993. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. J. Neurophysiol. 69, 1091–1117. https://doi.org/10.1152/jn.1993.69.4.1091.

Eckhorn, R., Krause, F., Nelson, J.I., 1993. The RF-cinematogram. Biol. Cybern. 69, 37–55. https://doi.org/10.1007/BF00201407.

Euler, T., Hausselt, S.E., Margolis, D.J., Breuninger, T., Castell, X., Detwiler, P.B., Denk, W., 2009. Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. Pflugers Arch - Eur J Physiol 457, 1393–1414. https://doi.org/10.1007/s00424-008-0603-5.

Fiorani, M., Azzi, J.C.B., Soares, J.G.M., Gattass, R., 2014. Automatic mapping of visual cortex receptive fields: a fast and precise algorithm. J. Neurosci. Methods 221, 112–126. https://doi.org/10.1016/j.jneumeth.2013.09.012.

Frost, B.J., Wylie, D.R., Wang, Y.C., 1994. The analysis of motion in the visual systems of birds. In: Davies, M.N.O., Green, P.R. (Eds.), Perception and Motor Control in Birds: An Ecological Approach. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 248–269. https://doi.org/10.1007/978-3-642-75869-0_15.

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

Grienberger, C., Konnerth, A., 2012. Imaging calcium in neurons. Neuron 73, 862–885. https://doi.org/10.1016/j.neuron.2012.02.011.

Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiology 160, 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837.
Hunter, P.R., Lowe, A.S., Thompson, I.D., Meyer, M.P., 2013. Emergent properties of the optic tectum revealed by population analysis of direction and orientation selectivity. J. Neurosci. 33, 13940–13945. https://doi.org/10.1523/JNEUROSCI.1493-13.2013.
Joesch, M., Plett, J., Borst, A., Reiff, D.F., 2008. Response properties of motion-sensitive visual interneurons in the lobula plate of drosophila melanogaster. Curr. Biol. 18, 368–374. https://doi.org/10.1016/j.cub.2008.02.022.
Johnston, J., Ding, H., Seibel, S.H., Esposti, F., Lagnado, L., 2014. Rapid mapping of visual receptive fields by filtered back projection: application to multi-neuronal electrophysiology and imaging: Rapid mapping of visual receptive fields by filtered back projection. J. Phys. 592, 4839–4854. https://doi.org/10.1113/jphysiol.2014.276642.
Krapp, H.G., Hengstenberg, R., 1996. Estimation of self-motion by optic flow processing in single visual interneurons. Nature 384, 463. https://doi.org/10.1038/384463a0.
Krapp, H.G., Hengstenberg, R., 1997. A fast stimulus procedure to determine local receptive field properties of motion-sensitive visual interneurons. Vision Res. 37, 225–234. https://doi.org/10.1016/S0042-6989(96)00114-9.
Krapp, H.G., Hengstenberg, B., Hengstenberg, R., 1998. Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. J. Neurophysiol. 79, 1902–1917. https://doi.org/10.1152/jn.1998.79.4.1902.
Kubo, F., Hablitzel, B., Dal Maschio, M., Driever, W., Baier, H., Arrenberg, A.B., 2014. Functional architecture of an optic flow-responsive Area that drives horizontal eye movements in zebrafish. Neuron 81, 1344–1359. https://doi.org/10.1016/j.neuron.2014.02.043.
Kühn, N.K., Gollisch, T., 2019. Activity correlations between direction-selective retinal ganglion cells synergistically enhance motion decoding from complex visual scenes. Neuron. https://doi.org/10.1016/j.neuron.2019.01.003.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.
Marmarelis, P.Z., Naka, K., 1972. White-noise analysis of a neuron chain: an application of the Wiener theory. Science 175, 1276–1278.
Matsumoto, A., Tachibana, M., 2017. Rapid and coordinated processing of global motion images by local clusters of retinal ganglion cells. Proc. Jpn. Acad., Ser. B, Phys. Biol. Sci. 93, 234–249. https://doi.org/10.2183/pjab.93.015.
Naumann, E.A., Fitzgerald, J.E., Dunn, T.W., Rihel, J., Sompolinsky, H., Engert, F., 2016. From whole-brain data to functional circuit models: the zebrafish optomotor response. Cell 167, 947–960. https://doi.org/10.1016/j.cell.2016.10.019. e20.
Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.
Orger, M.B., Smear, M.C., Anstis, S.M., Baier, H., 2000. Perception of fourier and non-fourier motion by larval zebrafish. Nat. Neurosci. 3, 1128.
Paninski, L., 2003. Convergence properties of three spike-triggered analysis techniques. Network 14, 437–464.
Portugues, R., Feierstein, C.E., Engert, F., Orger, M.B., 2014. Whole-brain activity maps reveal stereotyped, distributed networks for visuomotor behavior. Neuron 81, 1328–1343. https://doi.org/10.1016/j.neuron.2014.01.019.
Qian, H., Zhu, Y., Ramsey, D.J., Chappell, R.L., Dowling, J.E., Ripps, H., 2005. Directional asymmetries in the optokinetic response of larval zebrafish (danio rerio). Zebrafish 2, 189–196. https://doi.org/10.1089/zeb.2005.2.189.
Reichardt, W., Rosenblith, W.A., 1961. Sensory Communication.
Reiser, M.B., Dickinson, M.H., 2008. A modular display system for insect behavioral neuroscience. J. Neurosci. Methods 167, 127–139. https://doi.org/10.1016/j.jneumeth.2007.07.019.
Ringach, D.L., Bredfeldt, C.E., Shapley, R.M., Hawken, M.J., 2002. Suppression of neural responses to nonoptimal stimuli correlates with tuning selectivity in Macaque V1. J. Neurophysiol. 87, 1018–1027. https://doi.org/10.1152/jn.00614.2001.
Ringach, D.L., Hawken, M.J., Shapley, R., 2003. Dynamics of orientation tuning in Macaque V1: the role of global and tuned suppression. J. Neurophysiol. 90, 342–352. https://doi.org/10.1152/jn.01018.2002.
Rinner, O., Rick, J.M., Neuhauss, S.C.F., 2005. Contrast sensitivity, spatial and temporal tuning of the larval zebrafish optokinetic response. Invest. Ophthalmol. Vis. Sci. 46, 137–142. https://doi.org/10.1167/iovs.04-0682.
Rust, N.C., Schwartz, O., Movshon, J.A., Simoncelli, E.P., 2005. Spatiotemporal elements of Macaque V1 receptive fields. Neuron 46, 945–956. https://doi.org/10.1016/j.neuron.2005.05.021.
Schwartz, O., Chichilnisky, E.J., Simoncelli, E.P., 2002. Characterizing neural gain control using spike-triggered covariance. In: Dietterich, T.G., Becker, S., Ghahramani, Z. (Eds.), Advances in Neural Information Processing Systems 14. MIT Press, pp. 269–276.
Schwartz, O., Pillow, J.W., Rust, N.C., Simoncelli, E.P., 2006. Spike-triggered neural characterization. J. Vision 6, 13. https://doi.org/10.1167/6.4.13.
Srinivasan, M.V., Jin, Z.F., Stange, G., Ibbotson, M.R., 1993. Vector white noise': a technique for mapping the motion receptive fields of direction-selective visual neurons. Biol. Cybern. 68, 199–207. https://doi.org/10.1007/BF00224852.
Theunissen, F.E., David, S.V., Singh, N.C., Hsu, A., Vinje, W.E., Gallant, J.L., 2001. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Network 12, 289–316. https://doi.org/10.1080/net.12.3.289.316.
van Santen, J.P., Sperling, G., 1985. Elaborated reichardt detectors. J. Opt. Soc. Am. A 2, 300–321.
Wang, C., Yao, H., 2011. Sensitivity of V1 neurons to direction of spectral motion. Cereb. Cortex 21, 964–973. https://doi.org/10.1093/cercor/bhq176.
Wang, K., Hinz, J., Haikala, V., Reiff, D.F., Arrenberg, A.B., 2019. Selective processing of all rotational and translational optic flow directions in the zebrafish pretectum and tectum. BMC Biol. 17, 29. https://doi.org/10.1186/s12915-019-0648-2.
Warren, W.H., Kay, B.A., Zosh, W.D., Duchon, A.P., Sahuc, S., 2001. Optic flow is used to control human walking. Nat. Neurosci. 4, 213. https://doi.org/10.1038/84054.
Wertz, A., Gaub, B., Plett, J., Haag, J., Borst, A., 2009a. Robust coding of ego-motion in descending neurons of the fly. J. Neurosci. 29, 14993–15000. https://doi.org/10.1523/JNEUROSCI.3786-09.2009.
Wertz, A., Haag, J., Borst, A., 2009b. Local and global motion preferences in descending neurons of the fly. J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol. 195, 1107–1120. https://doi.org/10.1007/s00359-009-0481-0.
Xiao, Q., Frost, B.J., 2013. Motion parallax processing in pigeon (columba livia) pretectal neurons. Eur. J. Neurosci. 37, 1103–1111. https://doi.org/10.1111/ejn.12115.

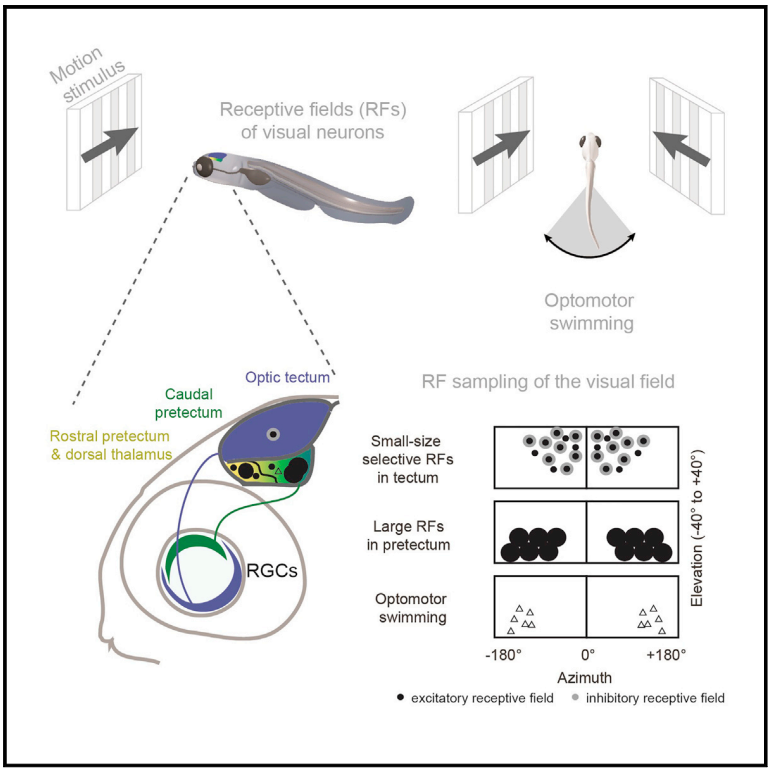
7.2 PARALLEL CHANNELS FOR MOTION FEATURE EXTRACTION IN THE PRETECTUM AND TECTUM OF LARVAL ZEBRAFISH

Article

Cell Reports

Parallel Channels for Motion Feature Extraction in the Pretectum and Tectum of Larval Zebrafish

Graphical Abstract



Highlights

- Optic tectum and pretectum sample visual space differently (upper nasal versus lower)
- Receptive fields are mainly small-size selective in tectum and large in pretectum
- Motion in the lower temporal visual field drives optomotor swimming
- The unequal sampling is likely an adaptation to hunting and optomotor behavior

Authors

Kun Wang, Julian Hinz, Yue Zhang, Tod R. Thiele, Aristides B. Arrenberg

Correspondence

aristides.arrenberg@uni-tuebingen.de

In Brief

Wang et al. show that the zebrafish optic tectum and pretectum extract visual motion features associated with hunting and optomotor behavior, respectively. The tectum preferably represents small stimuli in the upper nasal visual field, whereas the pretectum is biased toward larger stimuli in the lower visual field.

Wang et al., 2020, Cell Reports *30*, 442–453
January 14, 2020
https://doi.org/10.1016/j.celrep.2019.12.031

CellPress

OPEN ACCESS
CellPress

Cell Reports
Article

Parallel Channels for Motion Feature Extraction in the Pretectum and Tectum of Larval Zebrafish

Kun Wang,[1,2,4] Julian Hinz,[1,2,4,5] Yue Zhang,[1,2] Tod R. Thiele,[3] and Aristides B. Arrenberg[1,6,*]
[1]Werner Reichardt Centre for Integrative Neuroscience, Institute for Neurobiology, University of Tübingen, 72076 Tübingen, Germany
[2]Graduate Training Centre for Neuroscience, University of Tübingen, 72074 Tübingen, Germany
[3]Department of Biological Sciences, University of Toronto Scarborough, Toronto, ON M1C 1A4, Canada
[4]These authors contributed equally
[5]Present address: Friedrich Miescher Institute for Biomedical Research, 4058 Basel, Switzerland
[6]Lead Contact
*Correspondence: aristides.arrenberg@uni-tuebingen.de
https://doi.org/10.1016/j.celrep.2019.12.031

SUMMARY

Non-cortical visual areas in vertebrate brains extract relevant stimulus features, such as motion, object size, and location, to support diverse behavioral tasks. The optic tectum and pretectum, two primary visual areas in zebrafish, are involved in motion processing, and yet their differential neural representation of behaviorally relevant visual features is unclear. Here, we characterize receptive fields (RFs) of motion-sensitive neurons in the diencephalon and midbrain. We show that RFs of many pretectal neurons are large and sample the lower visual field, whereas RFs of tectal neurons are mostly small-size selective and sample the upper nasal visual field more densely. Furthermore, optomotor swimming can reliably be evoked by presenting forward motion in the lower temporal visual field alone, matching the lower visual field bias of the pretectum. Thus, tectum and pretectum extract different visual features from distinct regions of visual space, which is likely a result of their adaptations to hunting and optomotor behavior, respectively.

INTRODUCTION

Visual receptive fields (RFs) are specific regions in space where visual stimuli will alter the firing status of neurons (Spillmann, 2014). The ability of the visual system to extract useful information from the visual environment is directly related to the form, organization, and diversity of neuronal RFs within the vertebrate visual system. Task-relevant visual features are processed in parallel channels in the brain (Nassi and Callaway, 2009), starting in the retina (Baden et al., 2016).

The optic tectum and pretectum, two brain regions in the mes- and diencephalon, receive direct input from direction-selective retinal ganglion cells (Giolli et al., 2006; Hunter et al., 2013; Robles et al., 2014) and encode visual stimuli moving in different directions (Wang et al., 2019). These evolutionarily ancient structures share developmental origins with the superior colliculus (tectum) and part of the accessory optic system (AOS) (pretectum) in mammals. They support navigation and orienting behavior in zebrafish—which lack a visual cortex—already soon after hatching in 5-day-old larvae (Beck et al., 2004; Niell and Smith, 2005). Zebrafish are an important model organism for non-cortical vision research, but the division of feature extraction tasks between tectum and pretectum is still largely unknown. In particular, their roles in feature extraction in relation to behavioral tasks are crucial for a mechanistic understanding of sensorimotor transformations in zebrafish.

In the pretectal area and in the AOS of many vertebrates, neurons having large RFs with broad direction tuning curves are abundant (Britto et al., 1981; Grasse and Cynader, 1984; Masseck and Hoffmann, 2008; Simpson, 1984; Walley, 1967). Such large RFs should help the animal to distinguish wide-field optic flow from local motion and to estimate ego-motion. This computation is particularly important because many vertebrates use the outcome to stabilize gaze and body position (Portugues and Engert, 2009; Rinner et al., 2005). In larval zebrafish, both the optokinetic response (OKR) (Kubo et al., 2014) and the optomotor response (OMR) (Naumann et al., 2016) have been shown to rely on visual processing within the pretectum. In invertebrates, similar computations mediating OMR behavior were identified in the lobula plate, where horizontal system cells have large RFs with preferred directions matching the rotational optic flow around the yaw axis (Krapp et al., 2001). Additionally, it was shown that optogenetic manipulation of these neurons is sufficient to evoke yaw optomotor behaviors in fixed and tethered flies (Haikala et al., 2013; Busch et al., 2018). In zebrafish, the pretectum contains further anatomical sub-divisions (Yáñez et al., 2018), including structures involved in processing small visual stimuli during prey capture (Semmelhack et al., 2014; Muto et al., 2017), regions responsive to large-field motion stimuli (Kubo et al., 2014; Naumann et al., 2016), and a pretectal dopaminergic cluster providing input to the optic tectum (Tay et al., 2011). However, the RF properties of the pretectum at both the population and single neuron level are not known. It is also unclear how RF tuning within the pretectum may contribute to the production of visually mediated behaviors.

In contrast to the pretectum, RF sizes and locations for neurons within the zebrafish tectum have been described before (Niell and Smith, 2005; Sajovic and Levinthal, 1982; Bergmann et al., 2018; Preuss et al., 2014; Zhang et al., 2011). Tectal neurons have relatively small RFs, conforming to the idea that tectal

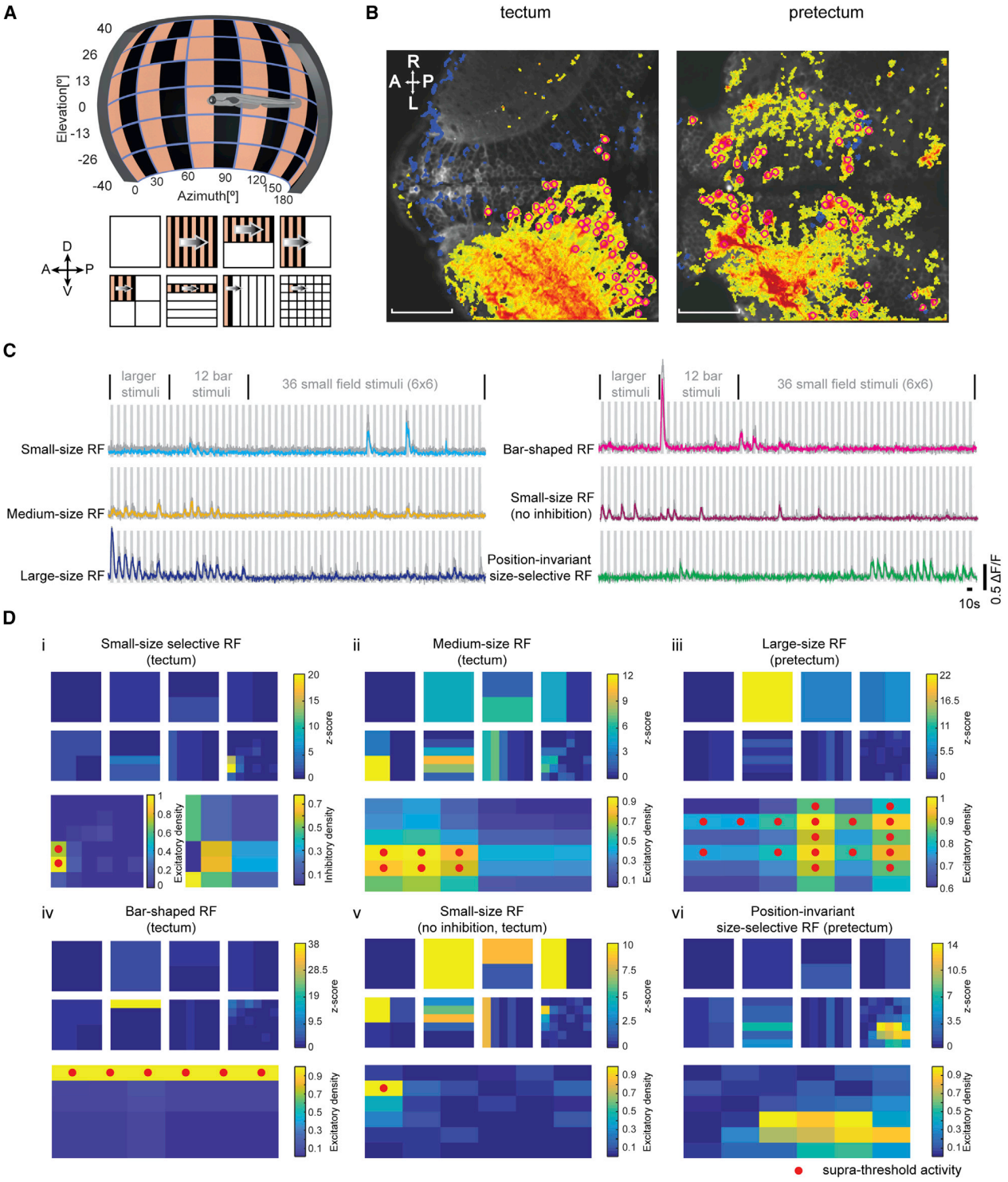


Figure 1. Different Receptive Field (RF) Types Were Identified Using Horizontally Moving Gratings

(A) Stimuli of naso-temporally moving gratings covering view fields of variable sizes were presented to the animal's right eye (n = 8 fish for naso-temporal and n = 2 fish for temporal-nasal motion). Motion stimuli consisted of whole-field (1 × 1, 180° × 80°, azimuth × elevation), half-field (2 × 1/2 and 1/2 × 2), quarter-field (2 × 2 × 1/4), bar (6 × 1/6 and 1/6 × 6), and small-field (6 × 6 × 1/36, each 30° × 13°) stimuli. The non-stimulated regions are shown in white for illustration purposes but contained a stationary grating. A snapshot during the small-field motion phase (depicted in the lower right) is shown on the display in the setup illustration (top).

(legend continued on next page)

neurons detect small-size moving objects in the view field and are needed for hunting behavior (Gahtan et al., 2005).

It is expected that each of the brain areas receiving input from the retina is adapted to the specific tasks and behaviors the animal executes in its environment. The behavioral relevance of a visual stimulus is influenced by the visual field location of the stimulus and will depend on the particular visually mediated behavior and the probability of observing such stimulus locations under natural conditions. For example, in the brain of macaque monkeys, it has been shown that RF properties in the superior colliculus, the homolog of the optic tectum, widely differ in the upper and lower visual view field, which likely represents adaptations to near space in the lower and far space in the upper visual field (Hafed and Chen, 2016). During hunting behavior, vertebrates typically keep the prey items, which are oftentimes small stimuli, in their nasal or frontal visual field, indicating that high visual acuity in this region of space is advantageous. Accordingly, both the primary visual cortex and superior colliculus show a magnification of foveal visual field regions (Grujic et al., 2018; Schwartz, 1980), i.e., more neurons are dedicated to representing these foveal locations than more peripheral locations. In the zebrafish retina, a region of heightened photoreceptor density (area centralis) has also been described (Schmitt and Dowling, 1999), corresponding to upper nasal visual field positions (Zimmermann et al., 2018). During prey capture, prey items need to be detected against the visual background, whereas for visual stabilization behaviors, animals need to detect their ego-motion by analyzing the global optic flow patterns resulting from the displacement of their bodies relative to the visual surround. To use brain resources efficiently, the reliable detection of optic flow directions is likely biased toward making use of the most informative visual field locations that occur in natural habitats and during behavior. The OMR is driven effectively by whole-field motion but—to our knowledge—there are no previous reports on particular visual field regions being preferably sampled by the animal to initiate OMR. Given the different roles of the optic tectum and the pretectum in hunting and stabilization behavior, respectively, it seems likely that these brain areas represent the visual field differently. It is unclear, however, whether the observed retinal anisotropies are relayed to primary visual areas in the zebrafish brain and whether magnifications of certain visual field locations exist in the tectum or pretectum of zebrafish. The characterization of such brain area-specific magnifications within the small vertebrate brain of larval zebrafish would advance our understanding of the efficient encoding of relevant information in the vertebrate brain and help to reveal the specific computations that brains have evolved to perform.

Here, we characterize the RF properties of tectal and pretectal motion-sensitive neurons using *in vivo* 2-photon calcium imaging of GCaMP5G transgenic animals and investigate their organization in visual and anatomical space. In addition, we investigate how the identified RFs match to the visual locations, which drive the OMR behavior. Our results reveal complementary roles of the optic tectum and pretectum to support behaviorally relevant motion feature extraction.

RESULTS

To estimate RF properties of pretectal and tectal neurons, we stimulated the right eye of immobilized larval zebrafish with a series of horizontally moving grating patterns of different sizes and locations (Figures 1A and S1; STAR Methods) and measured GCaMP5G calcium responses of neurons in the diencephalon and midbrain (Figure 1B). 1,926 motion-sensitive neurons that responded reliably during the three repetitions of the stimulus protocol were recorded in 10 animals. Neurons were divided into four functionally defined groups (Figures 1C, 1D, and S2; see STAR Methods for classification), based on the size and shape of their RFs: (1) small-size RFs, (2) medium-size RFs, (3) large-size RFs, and (4) bar-shaped RFs. RF sizes ranged from very small RFs (30° × 13°) to whole field (168° azimuth × 80° elevation). In neurons with smaller RFs, we oftentimes observed suppressive effects for larger motion stimuli (Figure 1Di), showing that these neurons were small-size selective. Small-size RF neurons without signs of inhibition were frequently encountered as well; even though the excitatory RF density (STAR Methods) was localized to a small patch in the visual field, the neurons were also responsive to whole-field stimuli (Figure 1Dv). Furthermore, some of the small-size and medium-size RF neurons each responded to small moving stimuli in a range of different visual field positions but did not respond to larger moving stimuli covering the same visual field locations, i.e., their responses were small-size selective and position invariant (Figure 1Dvi).

Pretectal RFs Are Larger Than Tectal RFs and They Are Less Often Size Selective

For each motion-sensitive neuron, we measured the location of the RF center in the visual field and the anatomical position of the soma in the brain (see STAR Methods). Based on morphological tectal borders visible in our brain volumes and previous anatomical annotations of the pretectum and tegmentum

(B) The average fluorescence of an example calcium-imaging time series recorded from tectum and pretectum (PT) is shown in gray. Motion-sensitive image pixels are shown in false color, with warm and cold colors corresponding to positive and negative motion phase correlation, respectively. Manually selected regions of interest (ROIs) are labeled with magenta circles. Scale bar, 50 μm; A, anterior; P, posterior; L, left; R, right.

(C) Example ΔF/F calcium responses of neurons with different RF sizes or shapes are shown. For each neuron, the colored trace corresponds to the median response across three repetitions (gray traces). The gray rectangular shades correspond to the 57 presented motion phases.

(D) RF maps for six example neurons corresponding to (C). Top: the eight squares correspond to the eight stimulus segments shown in (A), and each square corresponds to the stimulus arena surface (180° × 80°, azimuth × elevation). The calcium response is plotted as a *Z* score for each stimulus phase (for each ROI, the ΔF/F, subtracted by the average of the ΔF/F, divided by the standard deviation of the baseline ΔF/F). Bottom: by comparison of the activities evoked by spatially overlapping stimuli of different sizes, excitatory RF densities were calculated to measure the size of the RFs as the number of patches with suprathreshold activity (red dots). For cells with small-size excitatory RFs with maximal responses during the small-size stimulus phases (see STAR Methods), an inhibitory RF density was calculated to judge the extent of small-size selectivity (cell i). The anatomical location of each neuron is indicated (tectum or PT).

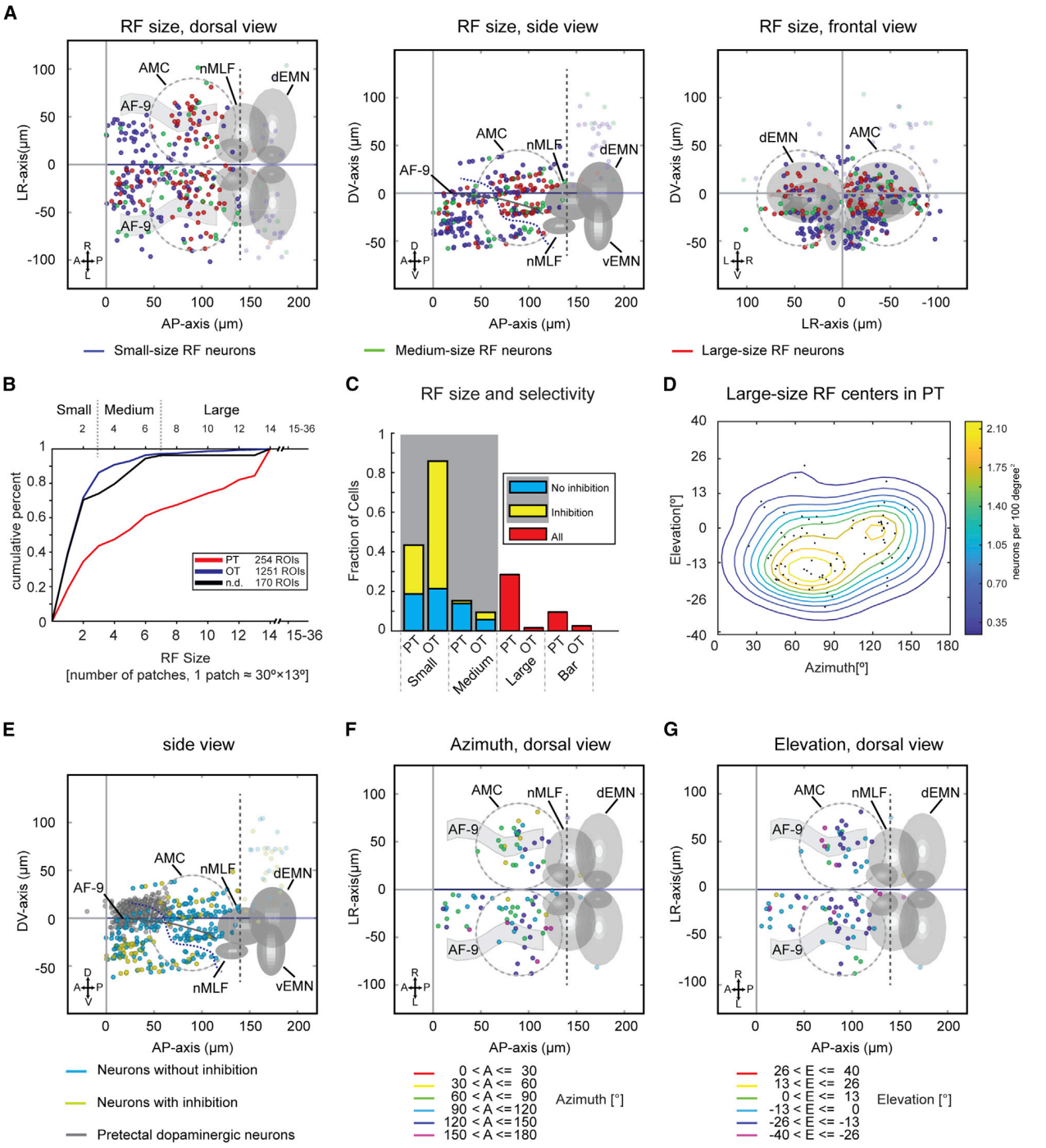


Figure 2. Pretectal RFs Are Large and Biased toward the Lower Visual Field

(A) Anatomical map of small-size (blue), medium-size (green), and large-size (red) RF neurons in the pretectal region. (n = 10 fish, 5 composite pretectal regions). Each colored dot represents a single neuron. Coordinates are defined as distances relative to the posterior commissure in the diencephalon (anterior-posterior axis and dorso-ventral axis) and midline (left-right axis). Many identified neurons were located within the previously annotated main pretectal cluster (AMC, dashed circle; Kubo et al., 2014) and additional neurons were located further rostrally (in the rostral PT and in the dorsal thalamus; Figures S6A–S6D). Neurons located ≥140 μm caudal to the posterior commissure (lighter colors) were located in the tegmentum and excluded from pretectal analysis (dashed gray line), whereas all rostral diencephalic neurons were included in the analysis.

(legend continued on next page)

(Kubo et al., 2014; Randlett et al., 2015), we identified tectal and tegmental neurons, as well as those located within the general pretectal area (Figure 2A). Neurons located in the tegmentum were excluded from further analysis. Within the general pretectal area, the caudal region corresponded to the pretectal anterior medial cluster (AMC) region described before (Kubo et al., 2014). In the recorded rostral diencephalic region, the dorsal part corresponded to the rostral pretectum (dorsal periventricular pretectal nucleus marked with dopaminergic neurons) and the ventral part to the dorsal thalamus annotated in the Z-Brain Atlas (Figures S6A–S6D) (Randlett et al., 2015; also see Discussion). However, the adult zebrafish brain contains several pretectal nuclei (Yáñez et al., 2018), and their exact identities and locations in the larval zebrafish brain still need to be resolved. We, therefore, included all diencephalic neurons in the analysis of the general pretectal area in this study. We investigated the topography of pretectal and tectal neurons as well as their sampling of the visual field. On average, we identified 205 ± 6 motion-sensitive neurons per tectum and 59 ± 13 neurons per pretectal region (n = 10 fish, corresponding to 5 complete composite brains sampled in 10-μm steps).

Pretectal neurons have larger excitatory RFs than those of tectal neurons (Figures 2B and 2C). Within the general pretectal area, 30% (88/295) of the motion-sensitive neurons had a large-size RF, compared to less than 2% (19/1,251) of the neurons within the tectum (Figure 2C). Most motion-sensitive tectal neurons (86%; Figure 2C) had excitatory RFs smaller than 1,200 deg^2 in area (3 of our small stimulus patches, each covering 30° × 13° in azimuth and elevation), which is consistent with previous reports from other groups (Sajovic and Levinthal, 1982; Bergmann et al., 2018). Furthermore, similar to the findings in a previous report (Preuss et al., 2014), we found that 68% (853/1,251) of motion-sensitive neurons in the tectum were size selective, i.e., motion stimuli that were larger than the neuron's excitatory RF evoked lower calcium responses (Figures 1Di and 2C). In contrast, only 26% (77/295) of the motion-sensitive neurons in the general pretectal area were size selective (Figure 2C).

Large-Size RFs in the Caudal Pretectum Are Biased to the Lower Visual Field

The RF centers of 69% of the large-size RF pretectal neurons were located in the lower visual field, which represents a significant bias (Figures 2D and S1B; p = 0.0002, z test for one proportion). These pretectal large-size RF neurons were located almost symmetrically in both hemispheres of the caudal pretectum, with some neurons in the rostral pretectum on the contralateral side (laterality index = −0.30; Figure 2A). The high number of ipsilateral neurons can—in part—be explained by reflections of the stimulus, which was revealed in an additional experiment in which the left eyes of the fish were blocked by a back foil (laterality index = −0.86; n = 6 fish, 6 pretecta; Figure S1C). In addition to the large-size RF neurons observed in the caudal pretectum, many neurons responsive to small moving stimuli were also identified (Figures 1Cvi, 2A, S6E, and S6F). These small-size RF neurons were most frequent in the rostral diencephalic region, which corresponds to the rostral pretectum and the dorsal thalamus (see Discussion; Figures S6A–S6D).

Because functionally identified neurons in the rostral diencephalon segregated from those in the caudal pretectum through an anatomical gap containing only few motion-sensitive neurons (Figure 2A, side view), we defined a boundary based on this gap to separate these two anatomical clusters (dashed line in Figures 2A, S4A, and S4B, and STAR Methods). The rostral diencephalic region contained a higher proportion of small-size RFs (57%, 85/148) than the caudal pretectum (29%, 43/147; p < 0.001, z-test for two proportions), whereas the caudal pretectum contained a higher proportion of large-size RFs (43%, 63/147) than the rostral diencephalic region (14%, 21/148; p < 0.001, z-test for two proportions) (Figures 2A and S1C). Furthermore, the rostral region contained a higher proportion of small-size selective neurons (Figures 2E and S4B–S4D), whose activity was suppressed for larger stimuli (rostral region: 34% [51/148], caudal pretectum: 18% [26/147], n = 148 and 147, p < 0.001, z-test for two proportions, see STAR Methods). Rostral diencephalic neurons responsive to our smallest grating stimuli were also frequently (63% of the neurons, n = 6 fish) responsive to small horizontally moving dots of variable diameters (3 to 18 degrees in diameter), as we tested in a separate experiment (data not shown).

It is well established that soma positions of zebrafish tectal neurons are topographically arranged within the tectum and that the RF centers cover almost the whole visual field at the population level (Attardi and Sperry, 1963; Niell and Smith, 2005; Romano et al., 2015; Bergmann et al., 2018). However, in the pretectum, we did not observe a clear topographic distribution of large-size RF neurons (Figures 2F, 2G, S1D, and S1E).

(B) Analysis of RF size differences across PT/diencephalon and optic tectum (OT). The cumulative distribution of RF sizes is shown for the general pretectal region within the diencephalon (PT, red), the OT (blue), and neurons of undefined provenance (black). Note that the PT has a larger fraction of large-size RF cells than the tectum.

(C) For each brain region (PT and OT), the fractions of small-size, medium-size, large-size, and bar-shaped RFs are shown. For small- and medium-size RFs, the inhibitory surround was investigated, and neurons with such inhibition are plotted in yellow.

(D) Locations and density contour plot of RF centers of large-size RF pretectal neurons in the contralateral and ipsilateral hemispheres (n = 10 fish, 5 composite brains).

(E) Anatomical map of three types of neurons in the pretectal region: neurons with (yellow) and without (cyan) signs of inhibition and pretectal dopaminergic neurons (gray). The dopaminergic neurons served as a landmark for the dorsal periventricular pretectal nucleus in the rostral PT and were recorded in a separate experiment.

(F and G) Topographic maps of large-size RF neurons in the pretectal region. Each colored dot represents a single neuron with its RF center in the indicated azimuth (F) and elevation (G) range (n = 10 fish, 5 composite pretectal regions).

Abbreviations are as follows: dEMN, dorsal extraocular motor neuron; vEMN, ventral extraocular motor neuron; dEMN and vEMN, the trochlear and oculomotor nuclei; nMLF, nucleus of the medial longitudinal fasciculus; A, anterior; P, posterior; D, dorsal; V, ventral, L, left; R, right; AF, arborization field. The abbreviations are applicable to all anatomical maps in this study.

Tectal RFs Dominate the Upper Nasal View Field

As expected from the topographic retino-tectal projection of retinal ganglion cells, strong rostro-caudal, dorso-ventral, and medio-lateral topographical gradients of RF centers were found for small-size RF neurons in the optic tectum (Figures 3B, 3D, 3E, and S3A). Responding tectal neurons were mainly located in the contralateral hemisphere relative to the stimulated eye (laterality index = −0.88; Figures 3D and 3E; see STAR Methods). More tectal neurons with small-size RFs corresponding to the upper nasal visual field were found than corresponding to the lower temporal visual field (420 versus 55 neurons, upper nasal versus lower temporal view field, *Z* score test for proportions: $p < 0.001$; Figures 3A and 3C). When comparing the upper nasal view field (26° elevation, 30° azimuth) to the lateral view field (0° elevation, 90° azimuth), the sampling difference corresponded to a 1.6-fold magnification factor (Figure S3B). In our experiments, we identified about 4 tectal small-size RF neurons per 10° × 10° in the upper nasal field in each completely sampled fish (hypothetical sampling every 5 μm in the dorso-ventral direction, calculated based on our actually recorded 33.4 tectal motion-sensitive neurons per 60° × 27° in each of 5 composite brains sampled every 10 μm; see Figures S3B–S3D).

Zebrafish OMR Behavior Is Driven Best by Motion in the Lower Temporal Visual Field

Although it is known that during hunting behavior, prey stimuli are mostly located in nasal visual field locations in zebrafish (Bianco et al., 2011), the visual field regions that drive OMR and OKR behavior have not been identified. OMR behavior can be stably induced with whole-field forward motion projected from below or from the side in larval zebrafish (Severi et al., 2014; Thiele et al., 2014), and it had been assumed that large—or even whole-field—stimuli are necessary to drive stabilization behaviors. Given the uneven distribution of large-size RFs in the upper and lower visual fields in the pretectum (Figure 2D), we wanted to test whether OMR is mainly driven by the forward motion located in the lower visual field, which would implicate the large-size RF neurons in mediating the OMR behavior. We, therefore, recorded the tail motion of larvae while forward-moving gratings of different sizes were presented in different visual field locations. Care was taken to always stimulate animals in a binocularly symmetric fashion to drive forward OMR instead of OMR turning behavior (Figures 4A, 4B, S5A, and S5B). The angle between the anterior-posterior body axis and the tail tip was traced to detect single tail beats and swim bouts (Figures 4A, 4D, and 4E). Bouts, consisting of a series of tail undulations beating symmetrically to both sides (forward OMR), were induced in response to whole-field forward-moving gratings (Video S1, S2, S3, and S4). In contrast, unsymmetrical unilateral turning swim beats to one side (turning OMR) were oftentimes evoked by whole-field rotating visual stimuli, although symmetrical beats (forward OMR) were also observed (Figure S5Fi). Whole-field and half-field visual stimuli covering the temporal or lower view field could evoke forward OMR robustly (Figure 4C).

To our surprise, forward OMR swim beats could be induced by stimuli as small as 45° × 20° (azimuth × elevation, the smallest size in our protocol) in the lower temporal view field of both eyes (Figure 4C). The OMR-evoking visual field locations were almost identical across all recorded animals (n = 6; Figure 4F). To test whether the OMRs evoked by small stimuli are stronger than what would be expected under the assumption that OMR drive was established by the sum of equal-sized motion inputs across the visual field, we normalized the evoked OMR tail-beat rate to the respective stimulus field size (analogous to the excitatory RF density estimation; see STAR Methods). The resulting visual field map of OMR drive (Figure 4G) shows the disproportionally large influence of moving stimuli in the lower temporal view field. We then compared OMR drive across different stimulus sizes (1 × 1, 2 × 1, … 4 × 4), always considering the visual field positions/stimulus phases that drove OMR best. This analysis revealed that the smallest stimulus area evoked responses, which were on average ~11 times stronger than expected by an equal integration of optic flow inputs across the visual field (Figure 4H).

DISCUSSION

Our study reveals the functional segregation of visual motion processing in parallel channels, each extracting different sets of motion features across the visual field. The optic tectum, which processes the motion of small visual stimuli, has a bias for upper nasal visual field locations. Within the diencephalon, wide-field optic flow is mainly processed in the caudal pretectum using large RFs that mainly sample the lower visual field, whereas small motion stimuli are mainly processed in rostral regions of the diencephalon. Furthermore, we show that animals observe mainly the lower temporal visual field for optomotor forward swimming.

These findings agree with the need to process small visual stimuli, e.g., during prey capture (Preuss et al., 2014; Bianco et al., 2011), and the need to assess wide-field motion to inform stabilization behaviors (Kubo et al., 2014). The data support a circuit model in which these two distinct tasks are processed independently by multiple channels in different brain areas.

Caudal and Rostral Diencephalic Regions Are Biased toward the Encoding of Large-Field Optic Flow and Small Stimuli, Respectively

Within the diencephalon, large-size RF neurons are mainly found in the caudal pretectal region, whereas small-size RF neurons are biased toward more rostral anatomical locations. The large-size RFs of pretectal neurons preferably sample the lower half of the visual field (Figures 2D and S1B), which fits with previous reports from pretectal neurons in dogfish (Masseck and Hoffmann, 2008). Due to our use of a half-cylindrical stimulus arena to present exclusively horizontally moving stimuli to the right eye, neurons with more complex, e.g., rotational, binocular, or vertical optic flow fields could not be described in this study (Kubo et al., 2014; Wang et al., 2019). These types of RF structures exist in visual neurons of other species (Krapp et al., 2001; Karmeier et al., 2003), and future studies are needed to identify them in zebrafish.

The adult zebrafish pretectum contains several nuclei distributed from the superficial to the periventricular regions, receives numerous retinal and tectal afferents, and projects to the optic

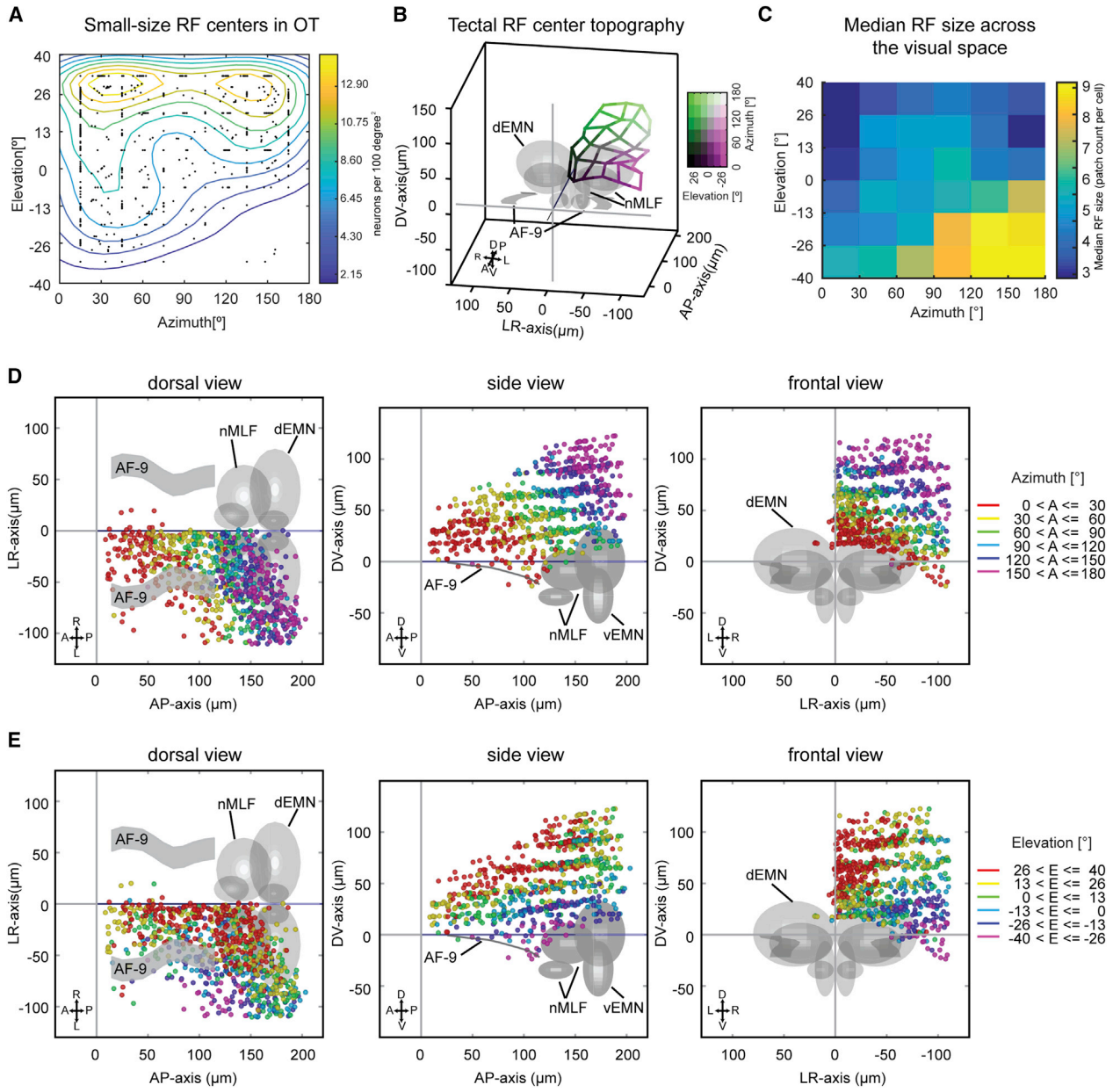


Figure 3. Tectal RFs Are Small and Biased toward the Upper Nasal Visual Field

(A) Visual field locations and density contour plot of RF centers of small-size RF tectal neurons (n = 10 fish, 5 composite tecta).

(B) Anatomical map of the tectal RF center topography. The average anatomical position of neurons with their RF centers in each of the 6 × 6 visual field bins was calculated and all 36 locations were connected by a grid to illustrate the mapping of visual space (color legend in the upper right) corresponding to anatomical space in the tectum (Figure S3A). The yaw angle (10° right) and the pitch angle (20° down) were adjusted to allow optimal view of the anatomical topography.

(C) Median RF size across the visual space. For each patch, we calculated the median RF sizes of all neurons (in tectum and PT) whose excitatory RFs covered the patch in question. The animals sample the lower temporal visual field mainly with large-size RF neurons, whereas small-size neurons dominate in the upper nasal visual field.

(D and E) Topographic maps of tectal small-size RF neurons for azimuth (D) and elevation (E). Each colored dot represents a single neuron with its RF center in the corresponding azimuth range in (D). For example, all RF centers of the neurons in red are located between 0° azimuth (in front of the fish) and 30° azimuth on the nasal right side of the fish. In (E), each colored dot represents a single neuron with its RF center in the corresponding elevation range. For example, RF centers of the neurons in green are located slightly above the equator of the view filed (0° to 13° in elevation). n = 10 fish, 5 composite brains.

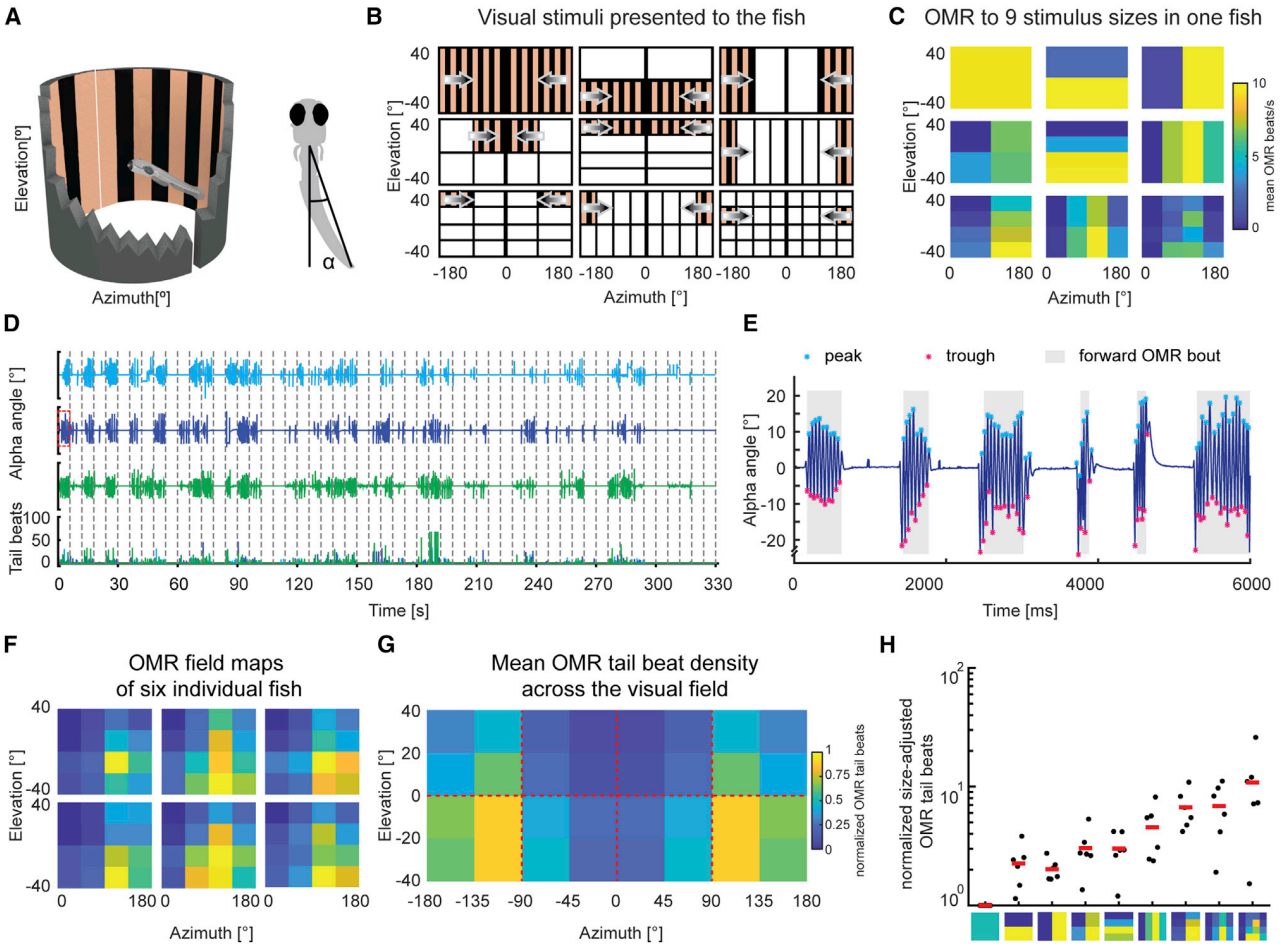


Figure 4. Zebrafish Optomotor Responses Are Driven by Motion from Below and in the Rear of the Animal

(A) Left: stimuli were presented binocularly to the animal with two half-cylindrical arenas while the behavior of the fish was recorded. A snapshot of the visual stimulus during the whole-field motion phase is shown in the setup illustration. Right: a schematic of a larval zebrafish performing OMR behavior. The indicated angle between the anterior-posterior body axis and the tail tip was measured to judge OMR performance.

(B) Motion stimuli of the OMR experiment consisted of whole-field (1x1, 180° azimuth x 80° elevation), half-field (2 × 1/2, 1/2 × 2), quarter-field (2 × 2 × 1/4), bar (4 × 1/4, 1/4 × 4), medium-field (4 × 2 × 1/8 and 2 × 4 × 1/8), and small-field (4 × 4 × 1/16, each 45° x 20°) stimuli (for each size group, only one motion patch is illustrated). Each stimulus was mirror-symmetric about the mid-sagittal plane and shown with two half cylindrical arenas from both sides. The control stimulus phases (not shown in B) consisted of two stationary whole-field phases, counterclockwise and clockwise rotational moving gratings or looming visual stimuli on either side (Figure S5B). The regions in which no stimulus movement was present are shown as white areas for illustration purposes but contained a stationary grating.

(C) Average number of tail undulations per second for one larva, induced by forward-moving gratings of 9 different sizes (rectangles represent the stimulus hemi-fields from B).

(D) The full behavioral session for one animal, showing induced OMR behavior by stimuli of different sizes and locations as indicated in (B). The three stimulus repetitions are shown in different colors (cyan, blue, and green). Individual stimulus phases are separated by the dashed gray lines, and each row corresponds to 55 concatenated time periods in which motion stimuli were presented. The stimulus pauses in-between motion phases were cropped out and lasted 69 s each. Measured tail-beat counts (bottom) during each trial were consistent across the three repetitions.

(E) Tail movements induced by whole-field forward-moving gratings (from trial No. 2 indicated by a red rectangle in D). The peaks and troughs of each swim bout within the 6-s recording are labeled with cyan and magenta asterisks, respectively. OMR swim beats are indicated by gray background shade.

(F) Visual field heatmaps of the forward-OMR tail-beat rate for six individual larval zebrafish (n = 6). The visual field density of OMR behavior was quantified analogous to the excitatory RF densities in Figure 1. The density was normalized according to the size of the stimulation area in each fish individually (see STAR Methods).

(G) Average heatmap of the OMR beat density for stimulation across different visual field coordinates (n = 6 fish).

(H) Maximum OMR beats for each of the 9 stimulus fields shown in (B) and (C) after normalization according to stimulus size (see STAR Methods).

tectum as well (Fernald and Shelton, 1985; Presson et al., 1985; Kastenhuber et al., 2010; Yáñez et al., 2018). The correspondence of these nuclei to functionally defined larval brain regions is not fully resolved (Kubo et al., 2014; Muto et al., 2017; Semmelhack et al., 2014; Arrenberg and Driever, 2013). In our study, the pretectal dopaminergic neurons, which are evolutionarily

conserved across most amniotes (Yamamoto and Vernier, 2011), were used as a landmark to indicate the location of the periventricular pretectal nucleus (Filippi et al., 2014). The location of caudal pretectal neurons in this study corresponds to the AMC, previously described by Kubo et al. (2014), as well as to the annotated pretectal brain volume in the Z-Brain Atlas (Figures S6A–S6C; Randlett et al., 2015). Recorded extra-tectal neurons located more than 140 μm caudal to the posterior commissure most likely belonged to the tegmentum and were, therefore, excluded from the pretectal analysis (see STAR Methods). The rostral diencephalic neurons described in this study cover rostral pretectal and dorsal thalamic brain regions (Videos S5, S6, and S7) (Rupp et al., 1996; Yáñez et al., 2018). Functional properties of two larval pretectal nuclei have previously been described in relation to prey capture (parvocellular and magnocellular superficial nuclei [PSp and PSm]); however, they are located more laterally than the bulk of our recorded motion-sensitive neurons in the rostral pretectum (Semmelhack et al., 2014; Muto et al., 2017). Morphologically, the identity, extent, and overlap of larval pretectal neuron populations, which give rise to each of the known adult pretectal nuclei, is not easily discernible (Arrenberg and Driever, 2013). Further anatomical studies are needed to link our functionally identified neurons to specific pretectal and thalamic brain nuclei and connectivity in the larval brain.

The anatomical and visual field locations, as well as neuron numbers, were consistent for large-size RF neurons in the pretectum across fish and experiments (Figures 2A, 2D, S1B, and S1C), whereas the responses of position-invariant RF neurons (Figure 1Dvi) appeared to be more variable. Notably, the preferred visual field locations, exact anatomical locations, and the number of identified position-invariant neurons within the diencephalon differed in the first and second experiment we performed (Figures S4E and S4F). Further work is needed to elucidate the specific anatomical distribution and response properties of position-invariant RF neurons.

The Tectum—Poised for Prey Capture

Our analysis of a large number of tectal neurons extends previous reports on tectal physiology (Niell and Smith, 2005; Sajovic and Levinthal, 1982; Zhang et al., 2011; Preuss et al., 2014). Our finding that the tectum mostly comprises relatively small excitatory RFs, which oftentimes are small-size selective, is in agreement with a role of the tectum in prey capture (Bianco et al., 2011; Gahtan et al., 2005). It is noteworthy that our stimulus protocol only allowed the measurement of RF sizes down to 30° horizontally and 13° vertically (corresponding to our smallest visual stimulus), which is much larger than the minimal RF size in a previous report (Preuss et al., 2014). Also, due to the limitation of 2-photon calcium imaging, we are not able to directly assess inhibition but assess it only indirectly by comparing the reduction in responses evoked by larger stimuli.

The topographic arrangement of the zebrafish tectal neurons is well established both for the anatomical retino-tectal projection (Baier et al., 1996; Trowe et al., 1996) and the functional RF mapping (Bergmann et al., 2018; Niell and Smith, 2005). However, precise measurements of dedicated tectal anatomical volumes had not been performed previously and are needed to build faithful models of zebrafish vision. We find that the upper nasal visual field (134 $\mu m^3/deg^2$) is magnified in the tectum by a factor of 1.6 relative to lateral (82 $\mu m^3/deg^2$) visual field locations (Figure S3), which is a relatively mild magnification in comparison to foveal magnification in the primate superior colliculus and visual cortex (Schwartz, 1980; Cowey and Rolls, 1974; Grujic et al., 2018). Given that the larval prey capture behavior depends on the optic tectum and larvae respond to the paramecia located in front and extending 60° temporalward of the fish (Bianco et al., 2011; Romano et al., 2015), a role of these tectal small-size RF neurons in prey capture seems likely. Because the density of small-size RF neurons is elevated in the upper nasal view field (Figures 3A and S3B), this would suggest that larval prey capture performance is best when the paramecia are located in front and slightly above the eyes and possibly the mouth. In this study, we have only investigated the RF distributions at a single developmental stage. It is possible that the reported tectal magnification of the upper nasal visual field is (in part) a result of the tectal developmental stage because new, initially non-functional neurons are added in the dorso-medial and caudo-lateral tectum (Boulanger-Weill et al., 2017; Recher et al., 2013).

The OMR Is Driven Most Strongly by the Lower Temporal Visual Field

Although it had been known from previous reports that motion stimuli presented from the bottom or from the side are effective in triggering OMR behavior (Thiele et al., 2014; Severi et al., 2014; Orger et al., 2008), it was unclear which parts of the visual field the animal preferentially responds to. We tested motion stimuli of variable size and position against a stationary background and show that for forward swimming OMR, the relevant region lies in the lower temporal view field of the fish (Figures 4F and 4G). We found that even motion stimuli as small as 45° × 20° (azimuth × elevation) can be effective OMR stimuli. This finding is in contrast to the concept that the OMR is a whole-field-induced behavior. Rather, it suggests that distinct parts of the visual field are sampled for body stabilization behaviors. This most likely reflects the ecological adaptions of zebrafish living in shallow waters (Engeszer et al., 2007) and the need to sample the most relevant parts of the visual field for different tasks (Zimmermann et al., 2018), such as putative high contrast textures within a river bed. Notably, zebrafish also sample different parts of their visual field for a related stabilization behavior, the OKR. The OKR of zebrafish is best driven by stimuli located laterally and slightly elevated (Dehmelt et al., 2019). Thus, each visually mediated behavior that has been investigated for visual field anisotropies (OMR, OKR, and prey capture behavior) preferentially samples information from different parts of visual space.

Because RGCs project directly to the optic tectum and pretectum (Robles et al., 2014), it seems likely that the small-size RFs without inhibition as well as the large-size RFs in the caudal pretectum are established by direct inputs from RGCs (Figure 5). Small-size-selective responses require inhibitory inputs, which could be calculated already within the retina or within the retino-recipient brain areas (Grama and Engert, 2012; Ramdya and Engert, 2008).

Previous work suggests that visual stabilization behaviors are driven by the caudal pretectum (Naumann et al., 2016; Wang

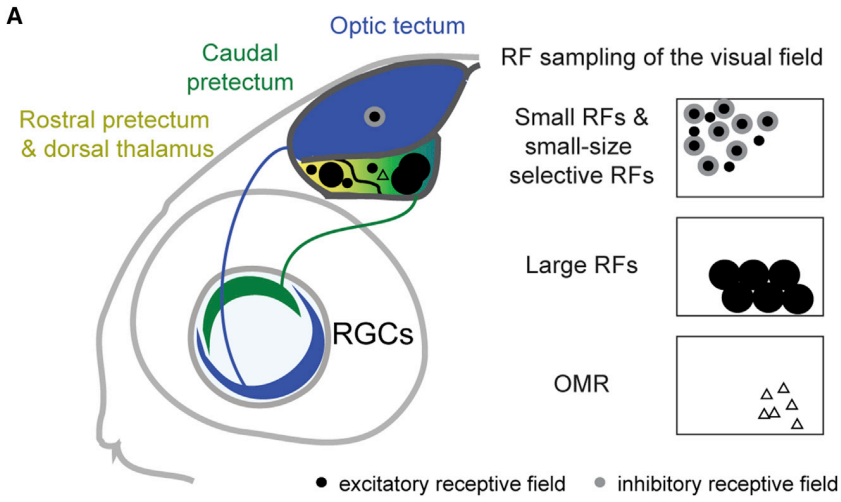


Figure 5. Illustration of the Major Anatomical Locations and Visual Field Preferences of the Characterized Functional Cell Types

The tectum mainly contains small-size-selective RFs, whereas a distribution of different RFs is observed in the PT and dorsal thalamus, ranging from small-size-selective (mainly in the rostral PT and dorsal thalamus) to large-size RFs (mainly in the caudal PT). Previous findings suggest that tectum and caudal PT receive direct retinal inputs (solid lines). To the right, the visual field distributions of RF centers are shown for each functional cell type. An additional visual field map is shown for OMR behavior, in which the regions driving OMR best are indicated. In the visual field maps, the horizontal axis (left-right) corresponds to the nasal-temporal spatial locations and the vertical axis to the upper and lower visual field locations.

et al., 2019; Kubo et al., 2014), which we show contains neurons of different RF sizes. Both whole-field and small-size optic flow stimuli evoke robust OMR tail beats, suggesting that both pretectal large-size and small-size RF neurons might form the basis for optic flow processing underlying the OMR. In agreement with this possibility, the RF centers of both pretectal large-size and small-size RFs match the bias of OMR drive to lower visual field locations (Figures S5E, S6E, and S6F). Further work is needed to identify the relative contributions of large-size RF and small-size RF pretectal neurons to OMR behavior.

In summary, we mapped the RFs of zebrafish tectal and pretectal neurons and demonstrated that both brain areas fulfill complementary roles for visual motion feature extraction. RFs of tectal neurons are predominantly small and size selective and have a strong bias in representing the upper nasal visual field (Figure 5). In contrast, caudal pretectal neurons have predominantly larger RFs with RF centers preferentially located in the lower visual field of the animal, which corresponds to the location of strongest OMR drive in the lower temporal visual field. Thus, each tectal and pretectal brain region extracts different motion stimulus features and samples distinct visual field regions. We speculate that this anisotropic visual field sampling in the tectum and pretectum could represent adaptations of zebrafish to feeding and stabilization behaviors, resulting in efficient usage of visual brain area volumes for the representation of behaviorally relevant stimulus features. Our study reveals the sensory layout of motion processing and, thus, constitutes an important advance for deriving a biologically faithful model of visuomotor transformations in zebrafish.

STAR★METHODS

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

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animal care and transgenic lines
- METHOD DETAILS
  - Animal preparation (tectal and pretectal imaging of neuronal somata)
  - 2-photon microscopy of somatic calcium responses
  - LED arena for visual stimulation (Figures 1, 2, 3, and 4)
  - Identification of regions of interest (imaging of neuronal somata)
  - Defining the borders of the tectum in our 3D datasets
  - Registration of 3D pretectal volume from Z Brain atlas (related to Figure S6)
  - Movies of 3D brain volumes with neurons shown inside
  - Monocular receptive field mapping
  - Inclusion criteria for somatic calcium responses
  - Setup for measuring the receptive field of OMR behavior
  - Receptive field mapping of OMR behavior
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.12.031.

ACKNOWLEDGMENTS

We thank Väinö Haikala and Dierk F. Reiff for help with the visual stimulus arena. The identification of the anatomical position of pretectal dopaminergic neurons was conducted in collaboration with Wolfgang Driever and Christian Altbürger, who provided anatomical z stacks of a double-transgenic dopaminergic line (unpublished data; see Fernandes et al., 2012 and Reinig et al., 2017). Furthermore, we thank Thomas Nieß (glassblower shop, University of Tübingen) and Klaus Vollmer (fine mechanics workshop, University Clinic Tübingen) for technical support, Prudenter-Agas (Hamburg, Germany) for generating illustrations, and Julianne Skinner for contributing to the analysis of RF center distribution. We thank Martin Meyer for feedback on a previous version of the manuscript. This work was funded by Deutsche Forschungsgemeinschaft (DFG) grants EXC307 (CIN-Werner Reichardt Centre

for Integrative Neuroscience) and INST 37/967-1 FUGG and Human Frontier Science Program (HFSP) Young Investigator grant RGY0079.

AUTHOR CONTRIBUTIONS

K.W. performed the experiments on pretectal and tectal somatic responses. K.W., J.H., and Y.Z. analyzed the data. A.B.A., K.W., J.H., and T.R.T. conceived the experiments and associated analysis protocols. K.W., J.H., and A.B.A. wrote the manuscript, with input from T.R.T.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 17, 2019
Revised: July 27, 2019
Accepted: December 9, 2019
Published: January 14, 2020

REFERENCES

Ahrens, M.B., Huang, K.H., Narayan, S., Mensh, B.D., and Engert, F. (2013a). Two-photon calcium imaging during fictive navigation in virtual environments. Front. Neural Circuits *7*, 104.

Ahrens, M.B., Orger, M.B., Robson, D.N., Li, J.M., and Keller, P.J. (2013b). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. Nat. Methods *10*, 413–420.

Arrenberg, A.B., and Driever, W. (2013). Integrating anatomy and function for zebrafish circuit analysis. Front. Neural Circuits *7*, 74.

Attardi, D.G., and Sperry, R.W. (1963). Preferential selection of central pathways by regenerating optic fibers. Exp. Neurol. *7*, 46–64.

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.

Baier, H., Klostermann, S., Trowe, T., Karlstrom, R.O., Nüsslein-Volhard, C., and Bonhoeffer, F. (1996). Genetic dissection of the retinotectal projection. Development *123*, 415–425.

Beck, J.C., Gilland, E., Tank, D.W., and Baker, R. (2004). Quantifying the ontogeny of optokinetic and vestibuloocular behaviors in zebrafish, medaka, and goldfish. J. Neurophysiol. *92*, 3546–3561.

Bergmann, K., Meza Santoscoy, P., Lygdas, K., Nikolaeva, Y., MacDonald, R.B., Cunliffe, V.T., and Nikolaev, A. (2018). Imaging Neuronal Activity in the Optic Tectum of Late Stage Larval Zebrafish. J. Dev. Biol. *6*, E6.

Bianco, I.H., Kampff, A.R., and Engert, F. (2011). Prey capture behavior evoked by simple visual stimuli in larval zebrafish. Front. Syst. Neurosci. *5*, 101.

Boulanger-Weill, J., Candat, V., Jouary, A., Romano, S.A., Perez-Schuster, V., and Sumbre, G. (2017). Functional Interactions between Newborn and Mature Neurons Leading to Integration into Established Neuronal Circuits. Curr. Biol. *27*, 1707–1720.e5.

Britto, L.R., Natal, C.L., and Marcondes, A.M. (1981). The accessory optic system in pigeons: receptive field properties of identified neurons. Brain Res. *206*, 149–154.

Busch, C., Borst, A., and Mauss, A.S. (2018). Bi-directional Control of Walking Behavior by Horizontal Optic Flow Sensors. Curr. Biol. *28*, 4037–4045.e5.

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.

Cowey, A., and Rolls, E.T. (1974). Human cortical magnification factor and its relation to visual acuity. Exp. Brain Res. *21*, 447–454.

Dehmelt, F.A., Meier, R., Hinz, J., Yoshimatsu, T., Simacek, C.A., Wang, K., Baden, T., and Arrenberg, A.B. (2019). Spherical arena reveals optokinetic response tuning to stimulus location, size and frequency across entire visual field of larval zebrafish. bioRxiv. https://doi.org/10.1101/754408.

Engeszer, R.E., Patterson, L.B., Rao, A.A., and Parichy, D.M. (2007). Zebrafish in the wild: a review of natural history and new notes from the field. Zebrafish *4*, 21–40.

Euler, T., Susanne, H., E., Margolis, D., J., Breuninger, T., Castell, X., Detwiler, P., B., and Denk, W. (2009). Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. Pflügers Archiv - European Journal of Physiology *457*, 1393–1414.

Fernald, R.D., and Shelton, L.C. (1985). The organization of the diencephalon and the pretectum in the cichlid fish, Haplochromis burtoni. J. Comp. Neurol. *238*, 202–217.

Fernandes, A.M., Fero, K., Arrenberg, A.B., Bergeron, S.A., Driever, W., and Burgess, H.A. (2012). Deep brain photoreceptors control light-seeking behavior in zebrafish larvae. Curr. Biol. *22*, 2042–2047.

Filippi, A., Mueller, T., and Driever, W. (2014). vglut2 and gad expression reveal distinct patterns of dual GABAergic versus glutamatergic cotransmitter phenotypes of dopaminergic and noradrenergic neurons in the zebrafish brain. J. Comp. Neurol. *522*, 2019–2037.

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.

Giolli, R.A., Blanks, R.H., and Lui, F. (2006). The accessory optic system: basic organization with an update on connectivity, neurochemistry, and function. Prog. Brain Res. *151*, 407–440.

Grama, A., and Engert, F. (2012). Direction selectivity in the larval zebrafish tectum is mediated by asymmetric inhibition. Front. Neural Circuits *6*, 59.

Grasse, K.L., and Cynader, M.S. (1984). Electrophysiology of lateral and dorsal terminal nuclei of the cat accessory optic system. J. Neurophysiol. *51*, 276–293.

Grujic, N., Brehm, N., Gloge, C., Zhuo, W., and Hafed, Z.M. (2018). Perisaccadic perceptual mislocalization is different for upward saccades. J. Neurophysiol. *120*, 3198–3216.

Hafed, Z.M., and Chen, C.Y. (2016). Sharper, Stronger, Faster Upper Visual Field Representation in Primate Superior Colliculus. Curr. Biol. *26*, 1647–1658.

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

Hunter, P.R., Lowe, A.S., Thompson, I.D., and Meyer, M.P. (2013). Emergent properties of the optic tectum revealed by population analysis of direction and orientation selectivity. J. Neurosci. *33*, 13940–13945.

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.

Karmeier, K., Krapp, H.G., and Egelhaaf, M. (2003). Robustness of the tuning of fly visual interneurons to rotatory optic flow. J. Neurophysiol. *90*, 1626–1634.

Kastenhuber, E., Kratochwil, C.F., Ryu, S., Schweitzer, J., and Driever, W. (2010). Genetic dissection of dopaminergic and noradrenergic contributions to catecholaminergic tracts in early larval zebrafish. J. Comp. Neurol. *518*, 439–458.

Krapp, H.G., Hengstenberg, R., and Egelhaaf, M. (2001). Binocular contributions to optic flow processing in the fly visual system. J. Neurophysiol. *85*, 724–734.

Kubo, F., Hablitzel, B., Dal Maschio, M., Driever, W., Baier, H., and Arrenberg, A.B. (2014). Functional architecture of an optic flow-responsive area that drives horizontal eye movements in zebrafish. Neuron *81*, 1344–1359.

Marques, J.C., Lackner, S., Felix, R., and Orger, M.B. (2018). Structure of the Zebrafish Locomotor Repertoire Revealed with Unsupervised Behavioral Clustering. Curr. Biol. *28*, 181–195.e5.

Masseck, O.A., and Hoffmann, K.P. (2008). Responses to moving visual stimuli in pretectal neurons of the small-spotted dogfish (*Scyliorhinus canicula*). J. Neurophysiol. *99*, 200–207.

Miri, A., Daie, K., Burdine, R.D., Aksay, E., and Tank, D.W. (2011). Regression-based identification of behavior-encoding neurons during large-scale optical imaging of neural activity at cellular resolution. J. Neurophysiol. *105*, 964–980.

Muto, A., Lal, P., Ailani, D., Abe, G., Itoh, M., and Kawakami, K. (2017). Activation of the hypothalamic feeding centre upon visual prey detection. Nat. Commun. *8*, 15029.

Nassi, J.J., and Callaway, E.M. (2009). Parallel processing strategies of the primate visual system. Nat. Rev. Neurosci. *10*, 360–372.

Naumann, E.A., Fitzgerald, J.E., Dunn, T.W., Rihel, J., Sompolinsky, H., and Engert, F. (2016). From Whole-Brain Data to Functional Circuit Models: The Zebrafish Optomotor Response. Cell *167*, 947–960.e20.

Niell, C.M., and Smith, S.J. (2005). Functional imaging reveals rapid development of visual response properties in the zebrafish tectum. Neuron *45*, 941–951.

Orger, M.B., Kampff, A.R., Severi, K.E., Bollmann, J.H., and Engert, F. (2008). Control of visually guided behavior by distinct populations of spinal projection neurons. Nat. Neurosci. *11*, 327–333.

Pologruto, T.A., Yasuda, R., and Svoboda, K. (2004). Monitoring neural activity and [Ca2+] with genetically encoded Ca2+ indicators. J. Neurosci. *24*, 9572–9579.

Portugues, R., and Engert, F. (2009). The neural basis of visual behaviors in the larval zebrafish. Curr. Opin. Neurobiol. *19*, 644–647.

Presson, J., Fernald, R.D., and Max, M. (1985). The organization of retinal projections to the diencephalon and pretectum in the cichlid fish, Haplochromis burtoni. J. Comp. Neurol. *235*, 360–374.

Preuss, S.J., Trivedi, C.A., vom Berg-Maurer, C.M., Ryu, S., and Bollmann, J.H. (2014). Classification of object size in retinotectal microcircuits. Curr. Biol. *24*, 2376–2385.

Ramdya, P., and Engert, F. (2008). Emergence of binocular functional properties in a monocular neural circuit. Nat. Neurosci. *11*, 1083–1090.

Randlett, O., Wee, C.L., Naumann, E.A., Nnaemeka, O., Schoppik, D., Fitzgerald, J.E., Portugues, R., Lacoste, A.M., Riegler, C., Engert, F., and Schier, A.F. (2015). Whole-brain activity mapping onto a zebrafish brain atlas. Nat. Methods *12*, 1039–1046.

Recher, G., Jouralet, J., Brombin, A., Heuzé, A., Mugniery, E., Hermel, J.M., Desnoulez, S., Savy, T., Herbomel, P., Bourrat, F., et al. (2013). Zebrafish midbrain slow-amplifying progenitors exhibit high levels of transcripts for nucleotide and ribosome biogenesis. Development *140*, 4860–4869.

Reinig, S., Driever, W., and Arrenberg, A.B. (2017). The Descending Diencephalic Dopamine System Is Tuned to Sensory Stimuli. Curr. Biol. *27*, 318–333.

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

Rinner, O., Rick, J.M., and Neuhauss, S.C. (2005). Contrast sensitivity, spatial and temporal tuning of the larval zebrafish optokinetic response. Invest. Ophthalmol. Vis. Sci. *46*, 137–142.

Robles, E., Laurell, E., and Baier, H. (2014). The retinal projectome reveals brain-area-specific visual representations generated by ganglion cell diversity. Curr. Biol. *24*, 2085–2096.

Romano, S.A., Pietri, T., Pérez-Schuster, V., Jouary, A., Haudrechy, M., and Sumbre, G. (2015). Spontaneous neuronal network dynamics reveal circuit's functional adaptations for behavior. Neuron *85*, 1070–1085.

Ronneberger, O., Liu, K., Rath, M., Rueβ, D., Mueller, T., Skibbe, H., Drayer, B., Schmidt, T., Filippi, A., Nitschke, R., et al. (2012). ViBE-Z: a framework for 3D virtual colocalization analysis in zebrafish larval brains. Nat. Methods *9*, 735–742.

Rupp, B., Wullimann, M.F., and Reichert, H. (1996). The zebrafish brain: a neuroanatomical comparison with the goldfish. Anat. Embryol. (Berl.) *194*, 187–203.

Sajovic, P., and Levinthal, C. (1982). Visual cells of zebrafish optic tectum: mapping with small spots. Neuroscience *7*, 2407–2426.

Schmitt, E.A., and Dowling, J.E. (1999). Early retinal development in the zebrafish, Danio rerio: light and electron microscopic analyses. J. Comp. Neurol. *404*, 515–536.

Schwartz, E.L. (1980). Computational anatomy and functional architecture of striate cortex: a spatial mapping approach to perceptual coding. Vision Res. *20*, 645–669.

Semmelhack, J.L., Donovan, J.C., Thiele, T.R., Kuehn, E., Laurell, E., and Baier, H. (2014). A dedicated visual pathway for prey detection in larval zebrafish. eLife *9*, 3.

Severi, K.E., Portugues, R., Marques, J.C., O'Malley, D.M., Orger, M.B., and Engert, F. (2014). Neural control and modulation of swimming speed in the larval zebrafish. Neuron *83*, 692–707.

Simpson, J.I. (1984). The accessory optic system. Annu. Rev. Neurosci. *7*, 13–41.

Spillmann, L. (2014). Receptive fields of visual neurons: the early years. Perception *43*, 1145–1176.

Tay, T.L., Ronneberger, O., Ryu, S., Nitschke, R., and Driever, W. (2011). Comprehensive catecholaminergic projectome analysis reveals single-neuron integration of zebrafish ascending and descending dopaminergic systems. Nat. Commun. *2*, 171.

Thiele, T.R., Donovan, J.C., and Baier, H. (2014). Descending control of swim posture by a midbrain nucleus in zebrafish. Neuron *83*, 679–691.

Trowe, T., Klostermann, S., Baier, H., Granato, M., Crawford, A.D., Grunewald, B., Hoffmann, H., Karlstrom, R.O., Meyer, S.U., Müller, B., et al. (1996). Mutations disrupting the ordering and topographic mapping of axons in the retinotectal projection of the zebrafish, Danio rerio. Development *123*, 439–450.

Walley, R.E. (1967). Receptive fields in the accessory optic system of the rabbit. Exp. Neurol. *17*, 27–43.

Wang, K., Hinz, J., Haikala, V., Reiff, D.F., and Arrenberg, A.B. (2019). Selective processing of all rotational and translational optic flow directions in the zebrafish pretectum and tectum. BMC Biol. *17*, 29.

Yamamoto, K., and Vernier, P. (2011). The evolution of dopamine systems in chordates. Front. Neuroanat. *5*, 21.

Yáñez, J., Suárez, T., Quelle, A., Folgueira, M., and Anadón, R. (2018). Neural connections of the pretectum in zebrafish (Danio rerio). J. Comp. Neurol. *526*, 1017–1040.

Zhang, M., Liu, Y., Wang, S.Z., Zhong, W., Liu, B.H., and Tao, H.W. (2011). Functional elimination of excitatory feedforward inputs underlies developmental refinement of visual receptive fields in zebrafish. J. Neurosci. *31*, 5460–5469.

Zimmermann, M.J.Y., Nevala, N.E., Yoshimatsu, T., Osorio, D., Nilsson, D.E., Berens, P., and Baden, T. (2018). Zebrafish Differentially Process Color across Visual Space to Match Natural Scenes. Curr. Biol. *28*, 2018–2032.e5.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
α-bungarotoxin	Sigma-Aldrich	T0195-.5MG
Poly-L-lysine solution	Sigma-Aldrich	P4832-50ML
Experimental Models: Organisms/Strains		
Zebrafish *Tg(HuC:GCaMP5G)a4598Tg*	Ahrens et al., 2013b	N/A
Zebrafish *Tg(th-E2A-QF2)m1512 x Tg(QUAS: EGFP)*	Driever lab (Freiburg University)	N/A
Software and Algorithms		
ImageJ/Fiji	NIH	https://fiji.sc
MATLAB R2010b, R2014b, R2015b	MathWorks	https://www.mathworks.com/products/matlab.html
LabVIEW 2015	National Instruments	http://www.ni.com/en-us/shop/labview.html
Mscan	Sutter Instrument	https://www.sutter.com/MICROSCOPES/mcs.html
Other		
Z-Brain Atlas	Randlett et al., 2015	https://engertlab.fas.harvard.edu/Z-Brain/

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Aristides Arrenberg (aristides.arrenberg@uni-tuebingen.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal care and transgenic lines

All animal procedures conformed to the institutional guidelines of the Universities of Tübingen and Freiburg and the local government (Regierungspräsidium Tübingen and Regierungspräsidium Freiburg, respectively). The transgenic zebrafish lines *Tg(HuC:GCaMP5G)a4598Tg* and the unpublished double transgenic line *Tg(th-E2A-QF2)m1512 x Tg(QUAS: EGFP)* (Fernandes et al., 2012; Reinig et al., 2017) were used in this study. Transgenic lines were kept in either a TL or TLN (nacre) background. Zebrafish larvae were raised in E3 medium until day 5 or 6 post-fertilization (dpf).

METHOD DETAILS

Animal preparation (tectal and pretectal imaging of neuronal somata)

At the day of experiments (5 or 6 dpf), larvae were transferred into a Petri dish and embedded in low melting agarose (E3 medium). The agarose surrounding the eyes was not removed as to minimize the range of possible eye movements. 6 animals received an injection of α-bungarotoxin into the caudal vein to paralyze them and prevent eye movements and motion artifacts. 4 animals were recorded without paralysis. In addition, 6 animals were recorded without paralysis in another independent experiment and exclusively used for the analyses presented in the Figures S1B, S1C, S4B, and S4F. The animals were then transferred and mounted in agarose on a glass triangle and the fish head protruded the point of the glass triangle, so that the eyes could see through the (agarose and) water clearly. The agarose surrounding the fish head was trimmed on the sides and in front of the animal to reduce the amount of surrounding agarose (Figure S1F). However, the eyes were still covered by agarose to minimize the range of possible eye movements. The glass triangle was held from the back by a 5 mm thin shaft which was fixed to an 8 cm diameter glass bulb (made by a glass blower) filled with E3 medium. The glass bulb resembled a consumer market light bulb (threading of the light bulb/glass bulb shaft at the back of the fish) and a 5 cm diameter hole was cut on the top of the spherical part to allow for approach of the microscope objective onto the fish (Figure S1F). From stimulus arena to the fish eye, the light traveled through air, glass (light hit glass roughly orthogonally in the spherical part as to minimize refraction of light rays), water, and finally agarose. The glass bulb was fixed with its shaft (15.5 mm diameter) to the metal holder which allowed for pitch and yaw adjustments (the glass shaft allowed for adjustments in roll). In the first batches of receptive field mapping (n = 6 larvae) experiments, we noticed a problem regarding reflections on the glass bulb on the opposite side

of stimulation, resulting in detected neurons on the ipsilateral side of (intended) stimulation in the tectum (which were not plotted in Figure 3). In the second set of recordings (4 larvae), we wrapped a piece of black, half-cylindrical aluminum foil around the objective. The black aluminum foil was then lowered beyond the eye contralateral to the stimulus to prevent this eye from seeing the reflections. In these recordings, only very few ipsilateral tectal neurons were detected (and were not excluded in the anatomical registration figures). Both sets of recordings were included in data analysis, because only few additional neurons were detected due to the reflections in the ipsilateral tectum, suggesting that the vast majority of the detected neurons of the contralateral side were detected due to the stimulus presented to the intended eye.

2-photon microscopy of somatic calcium responses

Calcium imaging was performed with a two-photon microscopy setup based on the MOM microscope (Sutter Instruments; Euler et al., 2009), using a Coherent Vision-S Ti-Sa laser and a 20x/1.0 Zeiss objective to image calcium signals in the transgenic fish line HuC:GCaMP5G (*Tg(elavl3:GCaMP5G)a4598*) (Ahrens et al., 2013b). Calcium time series were recorded at 2 frames per second, with an image size of 512 × 512 pixels and 2 x magnifications, at 920 nm, pre-pulse compensation set to 9756 fs^2. The midbrain and diencephalon were sampled from +60 μm below the landmark (posterior commissure) to −80 μm above the landmark. Optical slices were taken every 20 μm in the dorso-ventral direction in individual fish and across individual fish, and all dorso-ventral positions were recorded in 10 μm increments relative to the landmark (i.e., no recording at e.g., 5 or 15 μm below the landmark). Since we only recorded every 10 μm in dorso-ventral extent, more than twice as many neurons should have been detectable in the respective brain areas, had we sampled the brain areas at optimal spacing given the neuron soma diameter of ca. 5 μm. Where specified, error bars correspond to measures per completely imaged brain volume. Care was taken to record the same number of slices in each anatomical region. Two animals were used to image one complete brain volume (at 10 μm spacing). Due to the long recording times and positioning instability (likely resulting from the fish drifting within its agarose embedding), we corrected position drifts along the optical axis manually during the recording (mostly less than 4 μm per 30 minutes). Using the 20x objective and a magnification of 2x, our spatial resolution was 0.43 μm/pixel on the x axis (medial-lateral) and the y axis (anterior-posterior).

LED arena for visual stimulation (Figures 1, 2, 3, and 4)

Visual stimulation of zebrafish was conducted with a cylindrical LED arena consisting of 14336 LEDs (Kingbright TA08-81CGKWA): 2 (arena halves) x 8 (rows) x 14 (columns) x 64 (8x8 multiplexed LED matrix) LEDs. The caudal-most column of each arena half was removed without LEDs (i.e., 14, not 15 columns), since the space was needed for the glass bulb stage metal holder. Therefore, the caudal-most stimulus patches were slightly cropped (18° azimuth instead of 30° azimuth for the last patch). The arena covered −168° to + 168° in azimuth and −40° to 40° in elevation. A few degrees in angle of the dorsal field of view were likely blocked by the objective due to its access angle of 38.39° (< 40°), however the eyes were located ∼200 μm below the objective focus which should have resulted in a maximal viewing angle exceeding 38.39° (i.e., 39.2°). The LEDs emitted at 570 nm and an additional high-pass filter foil (LEE no. 779, article 595-1700-7790, castinfo.de, Hagen, Germany) and diffusion filter foil (LEE no. 252, article 595-1780-2520) were placed in front of the arena to optimize GCaMP signal detection and make the stimulus appear more homogeneous. This resulted in a yellow appearance of the stimulus. The LED arena was controlled as described previously (Joesch et al., 2008; Reiser and Dickinson, 2008). LEDs lit during fly-back time of the scanning mirrors.

Identification of regions of interest (imaging of neuronal somata)

For analysis of neuronal activity, a custom MATLAB script (MOM Load) identified regions based on their correlation to the stimulus and ROIs were manually drawn as described previously (Kubo et al., 2014; Miri et al., 2011). The 3-dimensional mapping of cell location was performed using custom written MATLAB scripts (Midbrain_Localizer and Cell_Viewer), which allowed to register the 2 dimensional recordings to a 3D z stack which was acquired after recording sessions (Kubo et al., 2014). See Figure S4 from Kubo et al. (2014) for an illustration of the 3-dimensional mapping procedure.

Defining the borders of the tectum in our 3D datasets

To distinguish pretectal from tectal neurons, we proceeded as follows: For each fish, the whole z stack – which was imaged from the top of tectum to deep ventral pretectum and dorsal thalamus – was resliced to generate a transverse view. On selected, regularly spaced transverse planes (more than ∼50 planes), the ventral border of the tectum was drawn: on each of these transverse plane (512 pixels from left to right, x dimension), a curve was drawn through the area devoid of neuronal somata or fluorescence that was ventrally adjacent to periventricular tectal area with densely packed, fluorescent somata. From each curve, 51 homogeneously distributed points were selected as key points with which a new boundary curve was generated by linear interpolation or three-term Gaussian fitting. Using this method, we obtained a boundary curve with 512 data points corresponding to the pixels in x/y dimensions (left-right and dorsal-ventral) for each transverse plane. In-between the annotated transverse planes, the 2D curves were interpolated to receive a surface that separated the tectum from the pretectum in all three dimensions (Wang et al., 2019). However, the boundaries between the caudal pretectum and adjacent brain areas in the posterior side (tegmentum) were not clearly visible in the GCaMP5G fish line. Referring to the AMC structure reported before (Kubo et al., 2014), and to the anatomical annotations of the

caudally adjacent tegmentum (Ronneberger et al., 2012; Randlett et al., 2015), the neurons below the tectal-pretectal boundary drawn above, which were located more than 140 μm caudally to the posterior commission, were excluded from the pretectum in the current study.

Registration of 3D pretectal volume from Z Brain atlas (related to Figure S6)

The pretectal region mask, obtained from datasets MaskDatabase.mat and Ref20131120pt14pl2.nrrd (Z Brain atlas, https://engertlab.fas.harvard.edu/Z-Brain/download) was used as the standard brain corresponding to the pretectal region mask.

(1) The brain regions corresponding to our recordings (mainly tectum and dorsal diencephalon) were cropped out from the standard brain (size 282 × 282, pixel x pixel; 1 pixel = 0.798 μm). The cropped image was resized to 512 × 512 (1 pixel = 0.43 μm) to match our image size and scale.
(2) The standard coordinate system which we use in our zebrafish brain was drawn in the standard brain of the Z Brain atlas (Figure S4 from Kubo et al., 2014). The relative locations of the pretectal region boundary to the point of origin (0, 0, 0) were calculated.
(3) The angle of the pretectal neuropil from the standard brain of the Z Brain atlas was measured with ImageJ. Compared to the standard pretectal neuropil angle of our own recordings (17.4°), the corresponding angle of the standard brain from the Z Brain atlas is pitched up, with a pretectal neuropil angle of 36.5°. Therefore, the pretectal region mask was rotated to register with our standard fish brain.
(4) The pretectal volume was plotted with MATLAB with the rotated data.

Movies of 3D brain volumes with neurons shown inside

The cell locations (e.g., the colored balls in Figure 2A) were registered to the 3D matrix which corresponds to a z stack of one larval brain, with the intersection point of the first two landmark lines as coordinate origin (0, 0, 0). In the 3D matrix, 3D spheres representing the neurons were plotted and the new 3D matrix was saved as an image series. Then the new image series and the standard zebrafish brain image series were merged with ImageJ. 3D brain volumes with neurons highlighted in different colors were generated with '3D Viewer' in ImageJ.

Monocular receptive field mapping

In the monocular receptive field mapping experiments, we used horizontally moving gratings, 0.033 cycles/°, moving at 30°/s, as visual stimuli to induce the neural activities. 8 fish were recorded using naso-temporal motion, and 2 fish were recorded using temporal-nasal motion. The data was pooled because we didn't observe obvious difference in RF characteristics (RF size, RF centers). The additional recordings for Figures S1B, S4B, and S4F were performed using both naso-temporal and temporal-nasal motion for 6 fish (3 temporal-nasal and 3 naso-temporal, respectively). Three repetitions of the 57 stimulus phases were shown to the right eye of the fish using one half-cylindrical arena (Figure 1A). All the stimulus patterns were presented in the order depicted in Figure S1A. In each trial, every 4.8 s stimulus phase was preceded by a 4 s pause and followed by a 2 s pause with the same stationary visual stimulus pattern. At the beginning and the end of each repetition, we inserted 9 s pauses.

RF maps for individual neurons were calculated by a series of analysis steps. First, we filter the DFF fluorescence traces with a low pass wavelet decomposition [type Daubechies, MATLAB: wavedec(DFF,1,'db4')] and a sliding median filter (the median of three data points). Then deconvolution was performed to the filtered data with the decay time constant (tau) of GCaMP5G, 1.5 s. We calculated the mean of phase-averaged signal (MPAS, averaged over stimulus phase time) from the deconvolved traces. The baseline was defined as the MPAS of all the non-stimulus phases (i.e., without moving stimulus). The standard deviation (STD) of all phase-averaged signals was calculated for the non-motion phases. And the z-score was calculated using the equation:

$$z-score = (MPAS - mean(baseline))/STD(baseline)$$

We then calculated the median MPAS z-score (i.e., the median across the three repetitions of a stimulus phase of the average of all data points within one stimulus phase).

To determine the size of receptive fields (RFs) of individual cells, we defined 5 subclasses of responses: small-size receptive fields, medium-size receptive fields, large-size receptive fields, bar-shaped receptive fields and double-field receptive fields (containing two discrete excitatory patches in the visual field). Since our stimulation protocol didn't allow for precise mapping of receptive fields (also Gaussian fits for larger receptive fields were problematic), we turned to a broad classification of RF sizes. To this end we used the smallest stimulation field (30° in azimuth, 13° in elevation) as a calculation unit (i.e., 1 "patch"). After manual inspection of the receptive field locations, we arbitrarily set the thresholds for the 3 size categories as 3 or less active phases (small, SM), 7 or less active phases (medium, ME) and 8 or more active phases (large, L). Bar cells were classified as having active phases only in the full vertical or full horizontal axis, while "double fields" were classified if they had two peaks of activation that were at least 60° away from each other. Most double-field receptive fields likely resulted from experimental

artifacts, in which stimulus reflections can cause such double field RFs. Neurons with double-field RFs were therefore excluded from further analysis.

To classify cells according to their respective size criteria, we ran a 2-step process:

Step I

1. Calculate mean patch density (MPD)
   a. If the cell responded maximally during the small 6x6 stimulus phases, the MPD corresponded to the average normalized activity of the cell in the phases of the 6x6 stimuli.
   b. If the cell responded maximally for one of the larger stimulus phases, the MPD was calculated as the average normalized activity of the cell in the phases of the 6x6 stimuli that were covered by the field of maximum activation (e.g., the cell in Figure 1Div had its maximal activity during in the upper horizontal bar stimulus phase, so the mean patch density would correspond to the average activity of the 6 upper patches of the 6x6 stimulation phase).
2. Next, we set the threshold for classifying a part of the visual field as "active" as follows: **IF**
   a. MPD is smaller than 30%:
      i. MPD * 3
   b. MPD is larger than 30% and smaller than 40%:
      i. MPD * 2
   c. MPD is larger than 40%:
      i. 90%

By relating the activity during the small stimuli to the activity observed during the larger stimuli, these MPD thresholds helped to obtain a more accurate quantification of active patches of cells preferentially active during the small stimulation phases.

3. In *Step I,* cells were classified if following criteria were met:
   a. The maximum activity in the 6x6 stimulation phase exceeded Mean patch density * factor (see step I.2) **or**
   b. Maximum excitation in the 6x6 stimulation phase larger than 90%
4. To classify the number of active patches, we applied the threshold defined in **point I.2**

Cells were then either classified as ME or SM based on the number of active patches.

All cells that didn't meet the criteria from point I.3 were classified according to ***Step II***. First, we calculated a second metric, the mean excitatory density (MED). We calculated the MED by multiplying the calcium response magnitude for each motion phase (excluding the 36 smallest motion phases) with an area factor (full size x 1, half size x 2 …), resulting in a motion phase's calcium activity weighted by the visual field area in which the stimulus was moving. We call this parameter the "excitatory density" of the stimulated part of the visual field. A biologically plausible underlying cause for differences in excitatory density is the number of DS RGC inputs the cell receives from the portion of the visual field in which the stimulus moves. We then summed the excitatory densities from all larger stimulation fields together (1x1, 2x1, 1x2, 2x2, 1x6, 6x1) taking into account their spatial location in the visual field. This resulted in an excitatory density map of the complete visual field covered by the stimulus arena. In order to report a single number for the RF size, we then defined active phases as those having 75% or more of the normalized summed maximum activation. The difference between this threshold and the 90% threshold for SM cells is derived from comparing manual and automated classification methods.

This method favors smaller receptive fields, because the calcium indicator only shows disproportionally small fluorescence levels for low levels of calcium activity - i.e., it is non-linear - but it enables easy classification of cells size preferences with the given limitations of the calcium indicator (Chen et al., 2013; Pologruto et al., 2004).

The results from our analysis, which is based on thresholding and classification, fits well with both the results obtained from an automated approach using PCA and clustering (Figure S2), and results from manual classification of receptive field sizes.

To determine if cells were small-size selective, we compared the responses to small-size and larger-size stimuli from our stimulus protocol. If (i) a cell was assigned to the small size (SM) or median size (ME) category and was identified during one of the 36 (6x6) small-size stimulation phases (see above), and (ii) the cell showed its maximum activity during the 6x6 stimulation then this cell was classified as being size selective ('Inhibition').

To determine the relative reduction in activity (relative to the response if only the small excitatory receptive field is stimulated) and to visualize the spatial structure of the inhibitory receptive field ("inhibitory density"), we summed the relative reduction of activity (analogous to the above described excitatory density) for every patch belonging to the RF and removed the patches that were part of the RF from the inhibitory field.

Please note that some cells were assigned SM or ME status based on the excitatory density map (and not based on the 36 small-size stimulation phases). We assigned the "no inhibition" status to all of these cells, because the maximal activity was not found in any of the small-size stimulation phases. However, the RFs of these cells could still show some form of inhibition (e.g., during presentation of larger half-field stimulus phases), which was not characterized here. The second cell in Figure 1Dii (an ME cell) is an example for such cells without assigned "inhibition" status.

We estimated the RF centers in XY space as the center of mass of the normalized activity in the active phases (those with red dots in Figure 1D), i.e., both location and level of activity in active phases determined the position.

The median receptive field size across the visual field was derived by calculating (for every visual field position) the median RF size of receptive fields that covered the respective visual field location to visualize the distribution of RF sizes in the visual space.

LED arena light rays that traveled roughly through the center of the glass bulb (where the fish head was located as well) hit the glass bulb wall on the opposite side with an angle of incidence of 90°. This resulted in about 4% of reflected light (according to the Fresnel equations) and this light was visible to fish eye that was not intended to be stimulated. In the first 6 animals, we noticed an unexpectedly high number of detected ROIs in the optic tectum ipsilateral to the stimulation. The vast majority of these tectal ipsilateral ROIs had a reversed retino-tectal topography indicating that these ROIs were detected because of the reflected light. We decided to exclude these ROIs in the anatomical reconstruction. For the other 4 animals, we blocked the non-stimulated eye by placing a half-cylindrical piece of black aluminum foil around the objective and lowering it below the level of the eye. In these animals, a much smaller number of neurons were detected in the ipsilateral tectum, and their anatomical location corresponded to the expected retino-tectal topography. In the data analysis, the small-size RF neurons in the ipsilateral hemisphere recorded from the first 6 animals were excluded. The laterality index for tectal neurons was −0.88 for the 4 animals with one blocked eye, and −0.46 for the 6 animals in which the eyes were not blocked.

In Figure 3A (locations of RF centers in the visual field for tectal small-size RF cells), only very few cells (55 out of 1074), which have RFs centers in the lower temporal of the visual field, were present. While we are convinced that this finding represents an actual under-representation of such cells in the optic tectum, we would like to discuss two experimental caveats, which can explain the effect partially (but not fully). First, the fish were mounted on a glass triangle and care was taken to allow free view of the stimuli from the position of the eye lenses by pushing the larva toward the tip of the glass triangle (thereby reducing the positional stability of the recording). However, for some animals, a small portion of the lower temporal visual field might have been blocked by the sides of the glass triangle. However, this caveat should only have affected extreme lower-temporal receptive field positions (e.g., > 120° in azimuth and < −30° in elevation). Second, the anatomical positions of those tectal neurons having lower temporal receptive field centers lie close to the border with the pretectum. The pretectum-classified small-size RF cells in proximity of the tectum (shade red dots in Figure 2A) can fill the gap in the ventral-caudal visual field in Figure 3A only partially (just 12 additional cells for the region > 90° azimuth and < 0° elevation) when such pretectal neurons are plotted together with the tectal neurons.

To characterize the distributions of RF centers across visual space, the density of the RF centers in the visual space was calculated with 'ksdensity' function in MATLAB. Contour lines were plotted based on the density.

One gap between the rostral diencephalic and caudal pretectal neurons is quite obvious (Figure 2A). Moreover, in the rostral diencephalon, many neurons are small-size selective (RFs with signs of inhibition). Therefore, a boundary was manually defined between the rostral diencephalon and caudal pretectum along this gap containing only few motion-sensitive neurons (Figure 2A). In the data analysis of the rostral diencephalic neurons, the rostral diencephalon was defined as follows: dorso-ventral axis < 10 μm and anterior-posterior axis < 60 μm, or dorso-ventral axis < −30 μm and anterior-posterior axis < 110 μm. 140 μm caudal to the posterior commissure along the anterior-posterior axis was conceded as the caudal boundary of the pretectum with other brain areas. The bilaterally symmetric anatomical distribution of the neurons relative to the midline was measured using a laterality index, which was calculated as,

$$\left(neurons_{right} - neurons_{left}\right) / \left(neurons_{left} + neurons_{right}\right).$$

Inclusion criteria for somatic calcium responses

We calculated the Pearson's linear correlation coefficients between the stimulus phase z-scores (see above) of the three stimulus protocol repetitions (for all 3 pairwise combinations) to characterize the reproducibility of stimulus-evoked calcium responses. In our further data analysis, we only kept the neurons for which all three correlation coefficients were higher than a certain threshold. The threshold was set between 0.65 and 0.75 to exclude around 30% neurons with low reproducibility of stimulus-evoked activity in the monocular receptive field mapping experiment (2004 out of 2995, 67% neurons were kept). We then performed signal-to-noise ratio (SNR) analysis to exclude neurons with unstable baseline. In the SNR analysis, a threshold of four was used on the z-score to detect positive neural responses. SNR was defined as the ratio of the average response of all the responsive phases to the standard deviation of the baseline. All neurons with SNR lower than a certain threshold were excluded. The threshold was set between 8 and 10 to exclude about 5% remaining neurons with low SNR. We kept about 96% (1926 out of 2004 neurons) in the monocular receptive field mapping experiment for further analysis.

Setup for measuring the receptive field of OMR behavior

The visual stimuli were presented binocularly with a 336° (from −168° to 168° in azimuth) surround LED arena (two half-cylindrical arenas, Figure 4A) from both sides of the fish, i.e., all stimulus phases (except the rotational and looming control stimuli) were mirror-symmetric across the midline. The image of the fish was reflected to a lens by a mirror positioned around 5 cm above the fish (1 cm above the glass bulb). An infrared-sensitive high speed camera (Model IDT iNdustrial Speed I, Integrated Design Tools Inc.) with an IR bandpass filter (ET 850/40, CHROMA) recorded (250 Hz) the behaviors of the fish through the lens (diameter,

25.4 mm; focal length, 100mm; THORLABS, LB1676) mentioned above. Since the mirror above the fish is tilted 45 degrees vertically, the light path from the mirror to the camera was horizontal. The fish was illuminated from below with a high power infrared LED light (850 nm, Conrad, Item No. 491248-62) positioned 1 cm below the glass bulb. The infrared light was mounted below a piece of milk glass and provided homogeneous background illumination around the fish (Figure S5A). The infrared LED light and the camera were triggered by the motion signal of the visual stimulus recorded by a LabVIEW DAQ box. In our experiment, the camera started recording about 300 ms after the motion phase onset of the visual stimuli. The infrared LED light was only on during the motion visual stimulus phases to reduce potential harm to the fish.

Receptive field mapping of OMR behavior

For each recording, a larval zebrafish (6 dpf) was transferred into a Petri dish and embedded in low melting agarose (E3 medium) on a glass triangle such that the fish body completely protruded the tip of the glass triangle. The agarose surrounding the larval tail was removed to free the tail. To reduce the amount of agarose surrounding the animal, agarose surrounding the fish was trimmed on the sides and in front. Only the agarose caudal to the animal attached to the triangular stage. Therefore, a large view field was accessible for the fish, and - more importantly - the contours of the beating tail could be imaged without optical obstruction from the glass triangle. The fish and the glass triangle were fixed to an 8 cm diameter glass bulb filled with E3 medium in the same way as described above (see Animal Preparation in the Method Details section). The animal was illuminated from below with a high power infrared LED light, which was triggered on 150 ms after the start of the grating motion and lasted for 10 s. During the first 6 s of each grating motion phase (300 ms delay), the animal was imaged with a high speed infrared camera at 250 frames per second. The video data were saved during the pauses and analyzed with custom written MATLAB code offline.

In the OMR behavioral test, forward moving gratings, 0.033 cycles/°, moving at 30°/s, were used as visual stimuli to induce the optomotor response. Three repetitions of the 55 stimulus phases were shown to both eyes of the fish using two half-cylindrical arenas. All stimulus patterns were presented in each of the three randomized orders depicted in Figure S5C for the repetitions. The animals were adapted to the stationary gratings for about half an hour before motion stimulation started. In each trial, every 9.75 s stimulus phase was preceded by an 18 s pause and followed by a 51 s pause with the same stationary visual stimulus pattern. Before presenting the 2nd and 3rd repetition of the 55 stimulus phases, the animals rested for about 1 hour.

The motion of the larval tail was traced using a custom MATLAB algorithm for image processing and the 'alpha angle' (the angle between the fish anterior-posterior body axis and the tail tip, see Figure 4A) was calculated. The single tail beats were detected by labeling the peaks and troughs of the alpha angle traces. In our analysis, only swim beats meeting the following two criteria were considered as forward OMR beats: (1) the tail-beat frequency (during individual swimming bouts) was higher than 25 Hz; (2) for each beat, the difference of the amplitudes of adjacent peaks and troughs divided by the sum of them was smaller than 0.3 and larger than −0.3 (symmetrical tail beats/forward swimming).

Two types of OMR tail beats were distinguished: symmetrical and unsymmetrical tail beats (Figures 4E and S5F). Unsymmetrical tail beats oftentimes occurred at the beginning and the end of OMR tail beat bouts (Figures 4E and S5F), similar to the case of freely swimming larval zebrafish (Marques et al., 2018). As expected, the larval zebrafish tried to turn in response to the rotating visual stimulus used as control in our protocol, with tails beating mainly unilaterally to the reverse direction of the rotation (Ahrens et al., 2013a). However, symmetrical tail beats were observed in the turning swimming bouts as well (Figure S5Fi).

In Figure 4 (and Figure S5), the shown tail-beat frequency does not correspond to the frequency reached during individual swim bouts, but instead corresponds to the average frequency during the stimulus phase, such that time periods without swim bouts lead to a reduction of tail-beat frequency.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical information calculated with MATLAB R2014b built-in functions is provided in each of the sections above. For statements of significance an alpha level of 0.05 was used unless stated otherwise.

The analyzed number of zebrafish and brains is indicated in the main text and figure legends. Error bars correspond to SEM unless stated otherwise.

DATA AND CODE AVAILABILITY

The scripts for data pre-processing and the pre-processed data are freely available from our G-Node repository (https://gin.g-node.org/Arrenberg_Lab/monocular_receptive_field_mapping). The original raw datasets have not been deposited at G-Node due to the large data size. All raw and processed data and software used to generate the figures will be made available upon request.

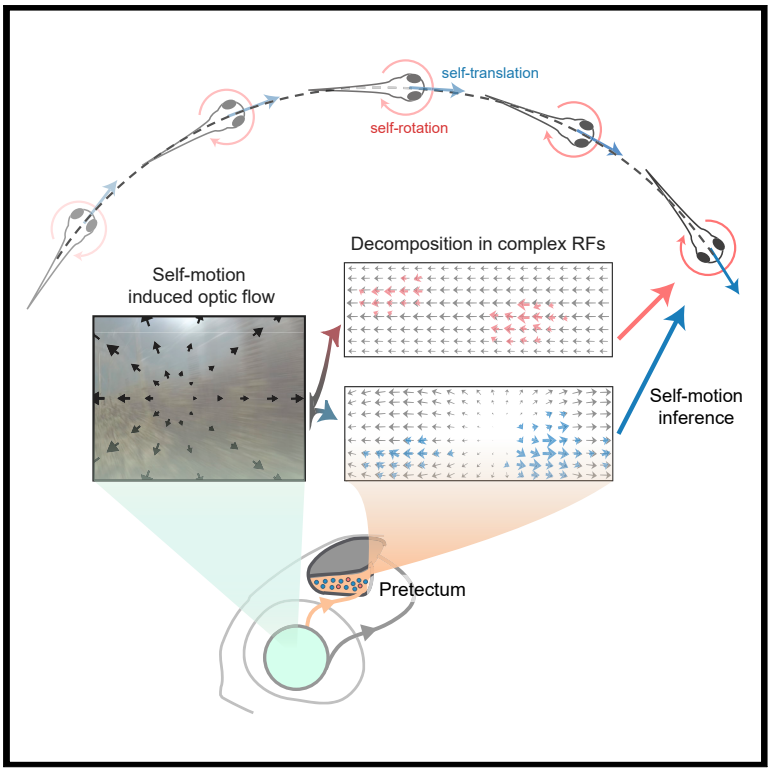
7.3 A ROBUST RECEPTIVE FIELD CODE FOR OPTIC FLOW DETECTION AND DECOMPOSITION DURING SELF-MOTION

Article

Current Biology

A robust receptive field code for optic flow detection and decomposition during self-motion

Graphical abstract



Highlights

- Motion noise responses reveal receptive fields (RFs) for optic flow processing
- These RFs serve as linear filters for self-motion extraction and decomposition
- A topographic map of self-translation direction is formed by neurons in pretectum
- The self-motion decomposition is consistently observed in animal behavior

Authors

Yue Zhang, Ruoyu Huang, Wiebke Nörenberg, Aristides B. Arrenberg

Correspondence

aristides.arrenberg@uni-tuebingen.de

In brief

Zhang et al. show how optic-flow-processing neurons extract self-motion information in zebrafish. A motion noise stimulus is used to discover a specialized group of receptive fields. These receptive fields may explain the robust detection and decomposition of self-translation and self-rotation observed in behavioral stabilization responses.

Zhang et al., 2022, Current Biology *32*, 1–12
June 6, 2022
https://doi.org/10.1016/j.cub.2022.04.048

CellPress

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

Current Biology

CellPress

Article

A robust receptive field code for optic flow detection and decomposition during self-motion

Yue Zhang,[1,2] Ruoyu Huang,[1] Wiebke Nörenberg,[1,2,3,4] and Aristides B. Arrenberg[1,5,6,*]
[1]Werner Reichardt Centre for Integrative Neuroscience, Institute of Neurobiology, University of Tübingen, 72076 Tübingen, Germany
[2]Graduate Training Centre for Neuroscience, University of Tübingen, 72076 Tübingen, Germany
[3]Present address: Humboldt-Universität zu Berlin, Faculty of Philosophy, Berlin School of Mind and Brain, Berlin, Germany
[4]Present address: Humboldt-Universität zu Berlin, Faculty of Life Sciences, Department of Psychology, Berlin, Germany
[5]Twitter: @ArrenbergLab
[6]Lead contact
*Correspondence: aristides.arrenberg@uni-tuebingen.de
https://doi.org/10.1016/j.cub.2022.04.048

SUMMARY

The perception of optic flow is essential for any visually guided behavior of a moving animal. To mechanistically predict behavior and understand the emergence of self-motion perception in vertebrate brains, it is essential to systematically characterize the motion receptive fields (RFs) of optic-flow-processing neurons. Here, we present the fine-scale RFs of thousands of motion-sensitive neurons studied in the diencephalon and the midbrain of zebrafish. We found neurons that serve as linear filters and robustly encode directional and speed information of translation-induced optic flow. These neurons are topographically arranged in pretectum according to translation direction. The unambiguous encoding of translation enables the decomposition of translational and rotational self-motion information from mixed optic flow. In behavioral experiments, we successfully demonstrated the predicted decomposition in the optokinetic and optomotor responses. Together, our study reveals the algorithm and the neural implementation for self-motion estimation in a vertebrate visual system.

INTRODUCTION

When you walk across a street, the image of the street is also shifted across your retina. This global retinal image shift induced by translational and/or rotational self-motion is known as optic flow.[1–3] Optic flow is used for self-motion estimation,[4–8] which is a critical task for all visually guided navigation, as well as for gaze and posture stabilization behaviors (Figure 1A).

The anatomical substrates associated with this task have been identified in a range of species. For instance, optic flow is processed in the medial superior temporal and parietal cortical areas in primates,[8–13] the nucleus lentiformis mesencephali in pigeons,[14] and the lobula plate in flies.[15] In zebrafish, the pretectal area processes optic flow and is necessary for body position and gaze stabilization in optomotor responses (OMRs) and optokinetic responses (OKRs), respectively.[16–19]

Knowing which neurons are involved does not answer how these neurons process optic flow and mediate stabilization behavior. For example, we cannot predict how an animal will respond to the optic flow induced by a banked turn, which contains both translation and rotation self-motion components. To answer this question, it is essential to reveal the underlying neural mechanism, or the algorithm, of optic flow processing. Theoretical studies have described several distinct biologically plausible algorithms.[20–22] One of these algorithms, the ''matched-filter'' model, suggests that motion receptive fields (RFs) of optic-flow-sensitive neurons form filters with similar structure to certain types of optic flow, i.e., each RF contains a vector field of locally preferred directions (PDs) that matches a particular self-motion axis. These matched filters may then form a mathematical basis of self-motion-induced optic flow and thus encode translation- or rotation-induced optic flow in a discriminable manner.[20,23]

In compact insect brains, systematic motion RF characterization has revealed the existence of such RF filters, which match certain types of translational or rotational optic flow (ROF).[15,24,25] However, there are many more neurons involved in optic flow processing in vertebrate brains than in insects. For example, the population of translation- and rotation-sensitive neurons in the pretectal area of larval zebrafish is at least 10 times larger than the population of horizontal system and vertical system neurons in blowfly.[15,16] It is thus unclear whether the matched-filter algorithm discovered in the compact fly brain is also employed by vertebrate brains. Also, with greater number of neurons comes greater difficulty in systematic RF characterization. Although calcium imaging enables recording of hundreds of neurons simultaneously, no efficient approaches had been available to characterize fine-grained motion RFs in vertebrate brains.

To overcome these limitations, we here made use of contiguous motion noise (CMN) stimuli and reverse correlation techniques developed recently for the systematic RF characterization of motion-sensitive neurons via calcium imaging.[26] We measured the fine-grained motion RFs of thousands of neurons in the pretectum and optic tectum (OT) of larval zebrafish. The identified RF structures reveal a neural mechanism for robust

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

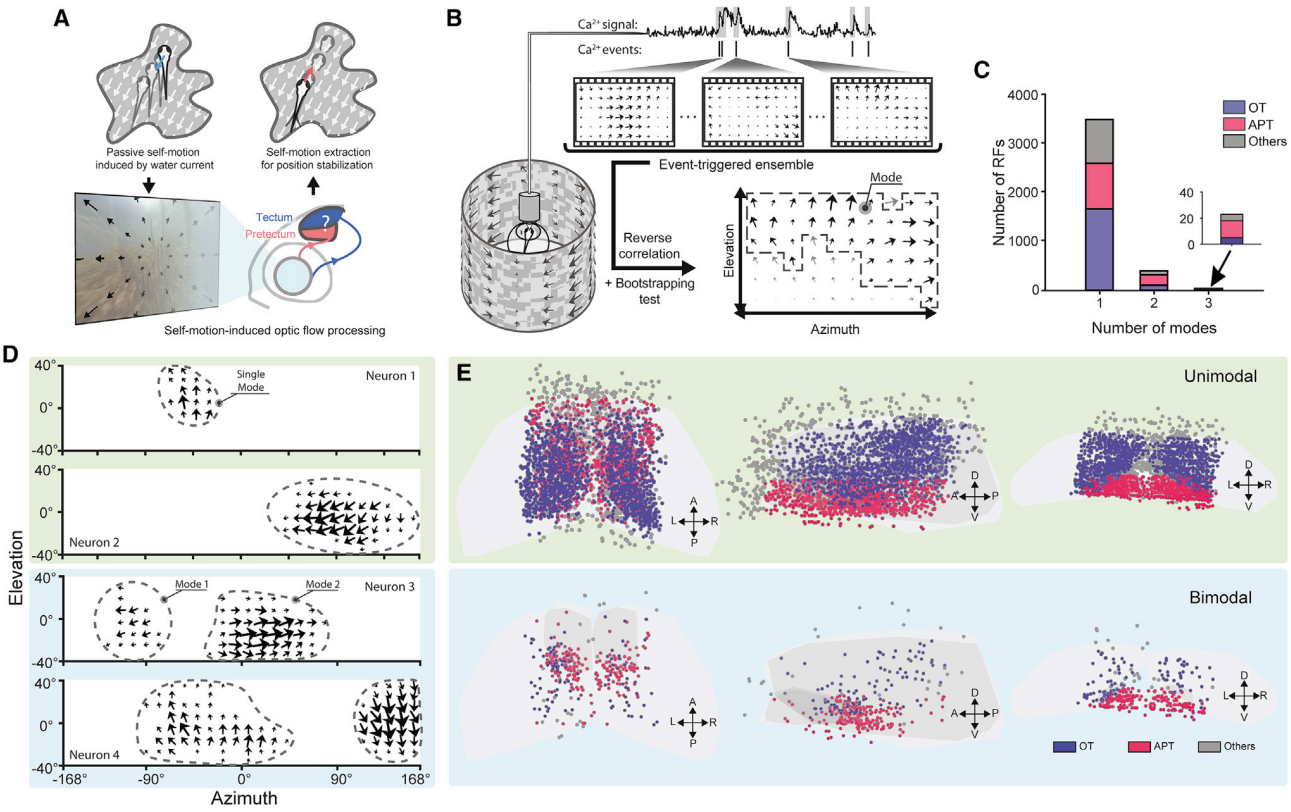


Figure 1. Characterization of unimodal receptive fields (RFs)

(A) Illustration of the challenge of extracting self-motion from optic flow for visually guided self-stabilization behavior.

(B) Illustration of the RF estimation method using a contiguous motion noise stimulus; see also STAR Methods and Figure S1.

(C) Histogram of RF mode number for neurons in the optic tectum (OT), area pretectalis (APT), and adjacent brain regions.

(D) Examples of two unimodal (green) and two bimodal motion RFs (blue). The direction and size of the arrows indicate the local preferred direction and the motion sensitivity within the RF; see also Figure S1.

(E) Anatomical location of neurons with unimodal RFs (n = 3,504, in 7 fish) and bimodal RFs (n = 398). The anterior-posterior (AP), dorsal-ventral (DV), and left-right (LR) axes are labeled.

optic flow encoding in the pretectum. Furthermore, we show that larval zebrafish decompose translation and rotation components in optokinetic and optomotor behavior as suggested by the underlying encoding mechanisms.

RESULTS

Identification of unimodal and bimodal RFs sensitive to motion stimuli

To estimate motion RFs, we used a full-surround cylindrical stimulus arena (Figure 1B) and recorded neuronal responses to motion stimuli in zebrafish larvae expressing the calcium indicator GCaMP6s. The stimulus consisted of a CMN pattern in which moving directions were spatiotemporally coherent. We then took the average of all stimulus frames that evoked calcium events in a neuron (reverse correlation) and applied bootstrapping statistics to identify which parts of these event-triggered averages (ETAs) can be trusted (Figures 1A and S1 and STAR Methods for details). The method has been described recently[26] and enabled us to identify the fine-scale RF structures of 3,926 motion-sensitive neurons from a total of 46,130 regions of interest (ROIs) recorded in the OT and the pretectal area of seven fish. The shape, span, location, and direction preference of these RFs were highly diverse, as illustrated by the estimated RF structures of the four example neurons in Figure 1D. This method enables inference of preferred motion directions in different parts of the visual field but is less suitable for estimating the precise diameter of the RF.[26]

Despite the high structural diversity, we found that most estimated RFs (99%) contained either one or two significant modes (Figure 1C), i.e., they could be classified into unimodal (or unipartite) and bimodal (bipartite) RFs. The "mode" here is defined as a patch of spatially connected RF components. Although both unimodal and bimodal RFs can be found in OT and the area pretectalis (APT), most unimodal RFs are located in the OT, and the bimodal RFs are predominantly located in the APT (Figures 1C and 1E).

The modes of bimodal RFs are located asymmetrically in the visual field and show patterned direction selectivity

Due to the fundamental structural difference between unimodal and bimodal RFs, we characterized them in different ways.

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

CellPress

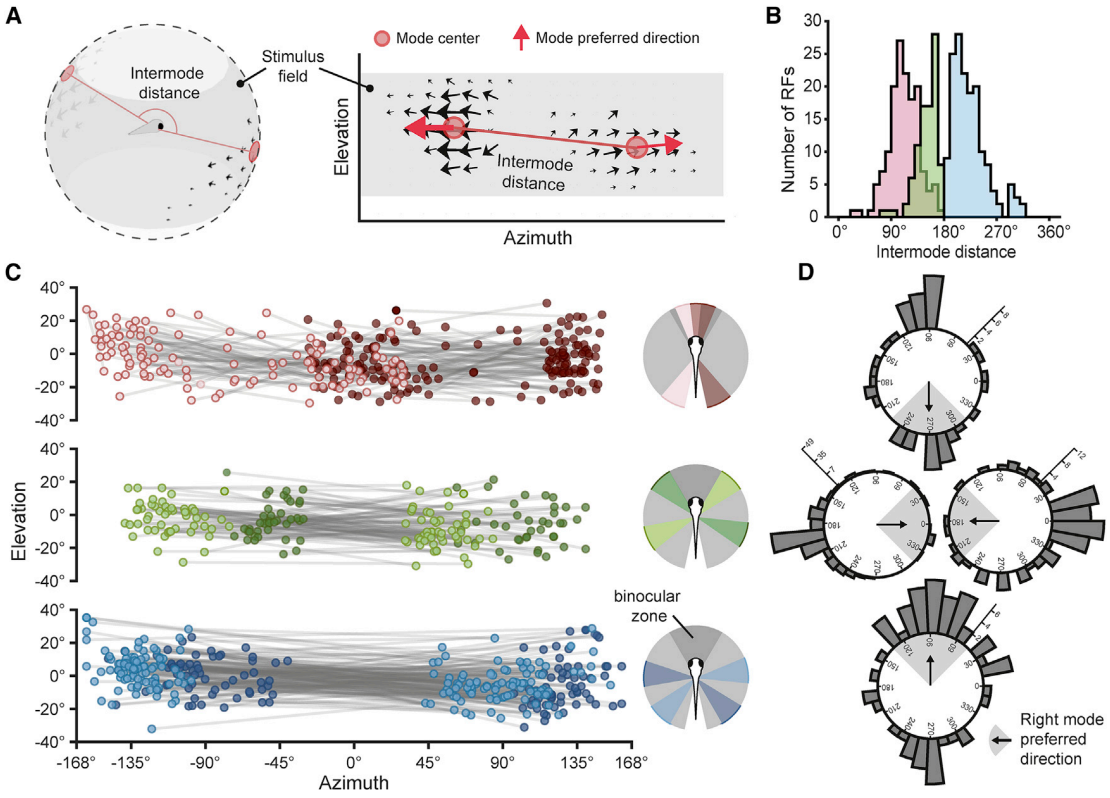


Figure 2. Spatial and directional organization of bimodal RFs across the visual field

(A) Illustration of a bimodal RF projected onto a sphere (left) and the corresponding 2D visual space representation (right). The "intermode distance" was defined as the distance between the two spatially separated components (modes) in the RF.

(B and C) Bimodal RFs grouped by their intermode distance and locations in visual field. Each pair of linked dots in (C) indicates the mode centers of a bimodal RF. The RFs tilted to the right hemi-field were colored darker for clarity. The approximate RF distributions within the visual field are summarized on the right.

(D) Pairing of preferred directions (PDs) for all bimodal RFs (i.e., mode pairs). The four polar histograms of the left mode PD are grouped by the PD of the right mode of each RF (indicated by the arrow and shade inside the polar histogram).

Each unimodal RF had an averaged PD that corresponded to either up, down, left, or right (Figure S2A), and the spatial RF distribution matched results of a previous study.[19]

For the bimodal RFs, the PDs and spatial centers were calculated separately for each mode. To measure the span of bimodal RFs, we calculated the distance between the two mode centers in visual space (intermode distance, illustrated in Figure 2A). The intermode distances ranged from ca. 60° to 315° in our dataset (Figure 2B), and intermode distances of 180° were scarce. The binocular overlap zone was smaller than 30° in our recordings[27] (cf. illustration in Figure 2C), as the eyes of the whole-mounted fish were in their normal resting position. Therefore, bimodal RFs spanning across two sides of the overlap zone were classified as binocular RFs. Based on this principle and the intermode distance, the bimodal RFs were divided into three groups: (1) putative monocular RFs receiving visual input from the nasal (−45° to +45° in azimuth) and temporal end (−168° to −90° or 90° to 168°) of a visual hemi-field (shown in red in Figures 2B and 2C), (2) binocular RFs with short intermode distance (average: ~150°, colored in green), and (3) binocular RFs with long intermode distance (average ~223°, colored in blue). For most bimodal RFs, the mode centers were distributed asymmetrically in relation to the anterior-posterior (AP) body axis as shown in Figure 2C.

Next, we compared the PDs within the two modes of each bimodal RF by plotting the distribution of PDs in the left mode grouped by their right mode PDs (Figure 2D). We found that the RFs preferring leftward (clockwise [CW] from above) motion with their right modes are mostly preferring the opposite direction (rightward motion) in their left mode, and vice versa. Among the neurons with vertical direction preferences in the right mode (31% of all bimodal neurons), 43% of them preferred motion in the opposite direction in the left mode (corresponding to rotation), while the other 57% showed similar direction preference (corresponding to lift translation; Figure 2D).

Complex bimodal RFs act as linear filters to encode optic flow directions

The directional preferences of the bimodal RFs (Figure 2D) implied these RFs may have similar structures to specific optic flow fields. For example, a bimodal RF that prefers leftward motion in its left mode and rightward motion in its right mode, is similar to the translational optic flow field (TOF) resulting from forward self-motion, and the RFs preferring upward and downward

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

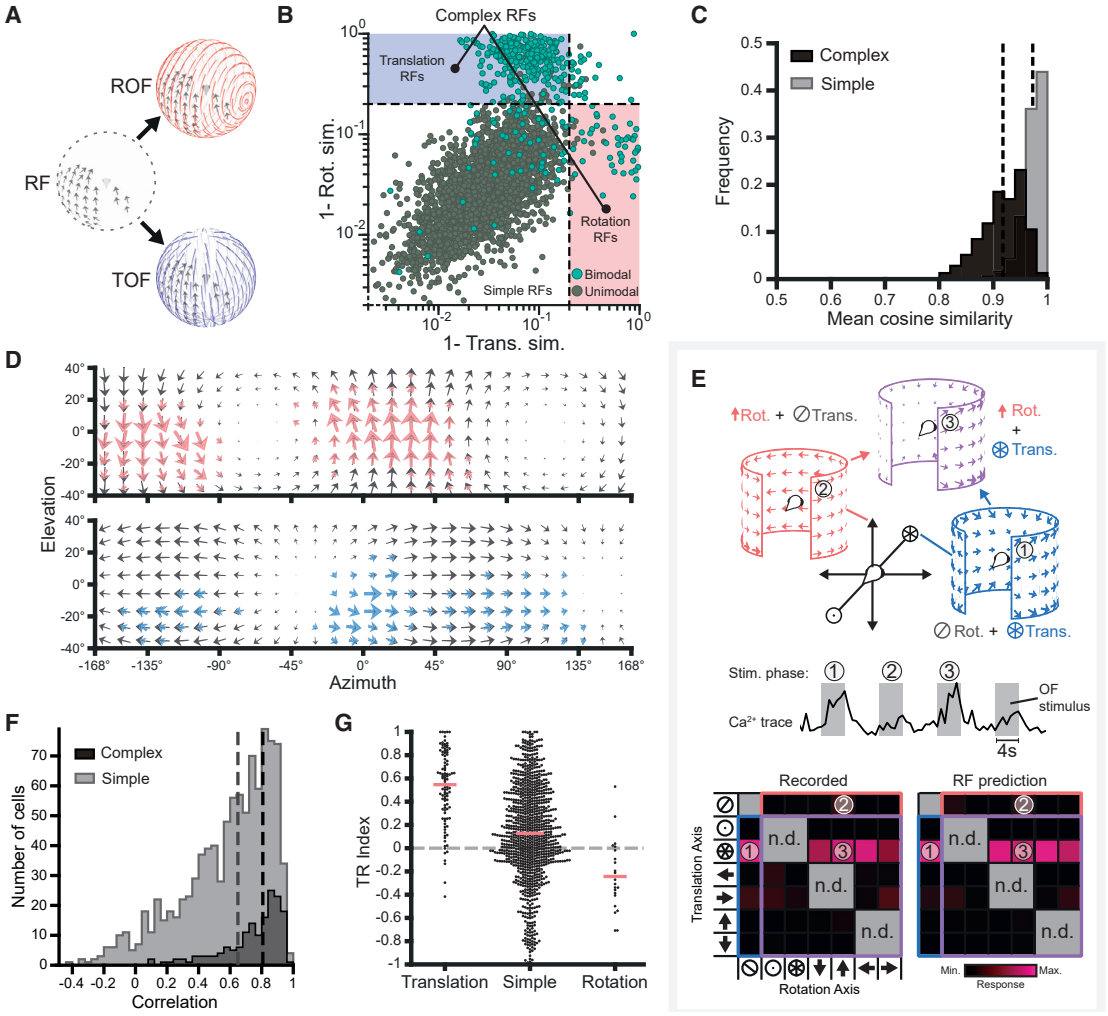


Figure 3. Prediction of optic flow responses in complex RF neurons with the matched-filter model

(A) The best-matching rotational (red) and translational (blue) optic flow field for an example motion RF (gray).

(B) Score of the match of each RF to the fitted rotation (rot.) and translation (trans.) optic flow field. Values below 0.2 (dashed lines) indicate good-matched-filter RFs; trans./rot. sim.: translation or rotation cosine similarity. RFs within blue and red regions were defined as "complex."

(C) Similarities of the complex and simple RFs to the fitted optic flow fields.

(D) Two examples of rotation (top) and translation (bottom) RFs (colored) and the matching optic flow fields (gray).

(E) The recorded maximum calcium responses of an example neuron to combinations of different optic flows (OFs, lower left) compared with the matched-filter model predictions (lower right). Three example OFs (1–3) are illustrated at the top and the neuron's corresponding calcium trace is shown in the middle row; n.d., not determined.

(F) The correlation of the recorded and predicted responses for complex and simple RFs.

(G) Bee swarm plot of the translation-rotation index (TR index; see also STAR Methods and Figure S2) for neurons with translation, simple, or rotation RFs. The dashed lines in (C) and (F) and the red horizontal lines in (G) indicate the median values of the corresponding distributions.

motion in their left and right modes, respectively, are similar to the ROF field resulting from roll. To quantify this similarity to optic flow fields, we calculated the average cosine similarity of all unimodal and bimodal RFs to their closest TOF and ROF (Figure 3A; STAR Methods section fitting optic flow fields to RFs). The similarity can vary from −1 to 1, where 1 means that the vector field structure of the RF is the same as the corresponding optic flow field, and 0 means that the RF is not similar to any optic flow field.

For most RFs (89%), both highly similar TOFs and ROFs existed (cf. dots within the bottom left quadrant in Figure 3B). These RFs were called "simple RFs" because most of them are unimodal RFs (98%) of relatively simple structure and small spatial extent, and thus, they can be interpreted as parts of TOFs and ROFs at the same time. Most of the other RFs were similar to either TOFs or ROFs only (cf. color-shaded areas in Figure 3B). These RFs were called "complex RFs" as they are mostly bimodal RFs (319 out of the total 389 complex RFs) containing heterogeneous vector field structures. In addition, most bimodal RFs (80%) are complex RFs. In these complex RFs, the 329 RFs that were more similar to TOFs were called

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

CellPress

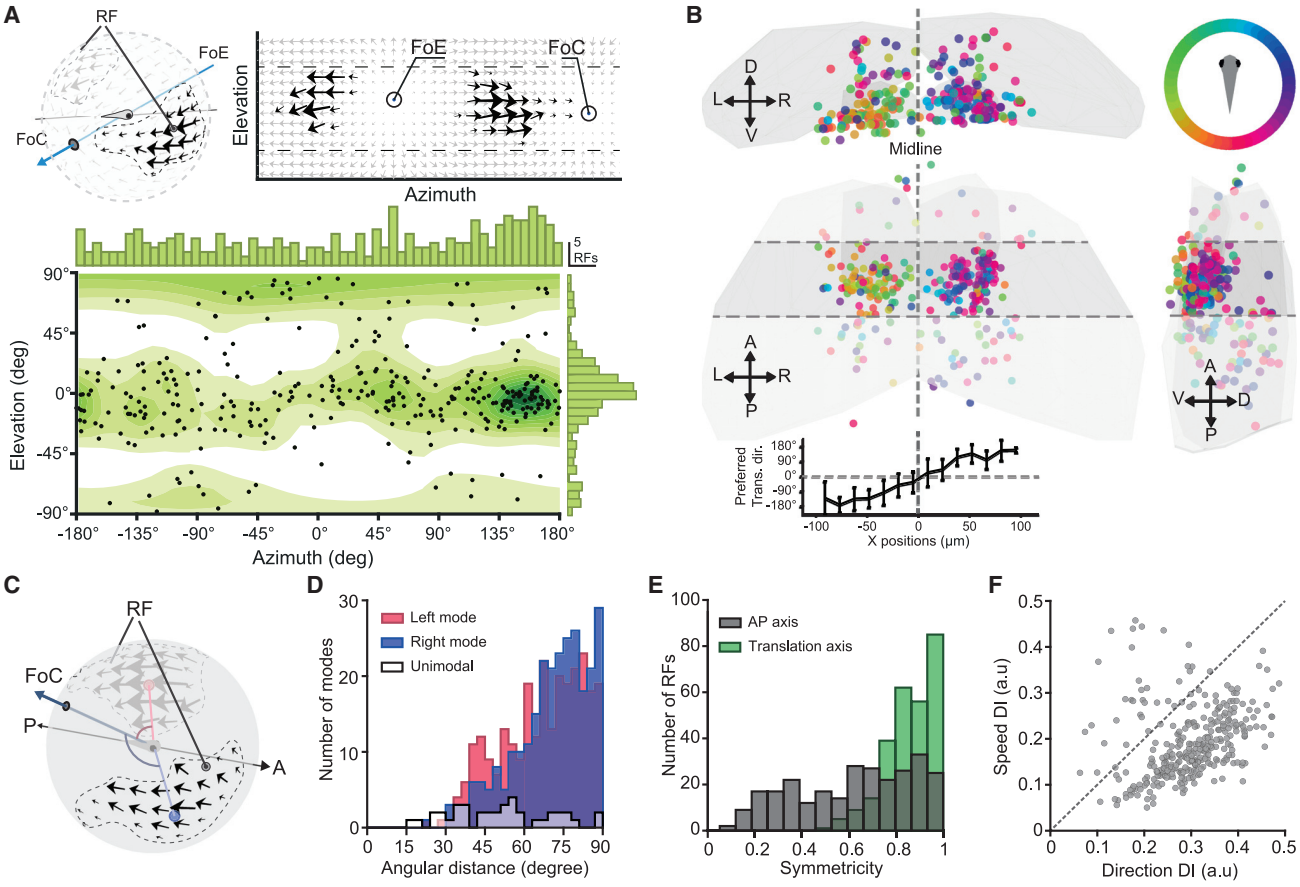


Figure 4. A topographic map of global translation direction in the pretectum

(A) Top: illustration of focus of expansion (FoE) and focus of contraction (FoC) of the fitted TOF field (gray arrows) of one translational RF (black arrows). Bottom: preferred translation direction distribution for all translation RFs (n = 329). Each dot corresponds to the preferred translation direction (FoC) of a neuron.

(B) Anatomical locations of neurons (n = 329) with translational RFs, colored according to preferred translation direction in the horizontal plane. The AMC pretectal region is contained within the horizontal dashed lines. The averaged preferred translation directions (trans. dir.) of the neurons binned by the location on the left-right axis (X position; bin size = 14.4 μm) are plotted at the bottom; the error bars indicate the SDs.

See also Videos S1 and S2.

(C) Illustration of the angular distance between mode centers and preferred translation direction (FoC) analyzed in (D). The red and blue arcs indicate the angular distance of left and right mode centers, respectively.

(D and E) The angular distance of mode centers in relation to the preferred translation direction (D) and the symmetry (E) of mode pairs relative to translation direction and body axis.

(F) The discrimination index (DI) for the tuning to translation speed (within the range of 7.5°/s–120°/s) and direction (0°–360°) of the 369 translation-sensitive neurons in the AMC of two fish.

See also Figure S4 and Video S4.

translation RFs, while the 60 RFs similar to ROFs were rotation RFs. Two flow field examples for translation and rotation RFs are shown in Figure 3D. Both simple and complex RFs can be well explained by the fitted ROF and/or TOF (Figure 3C): the medians of the cosine similarities for simple and complex RFs are 0.98 and 0.92, respectively.

This finding is consistent with the observation in flies, where the RF organization of direction-selective wide-field neurons were concluded to serve as matched filters indicating the presence of particular self-motions.[23] To test whether or not these complex RFs in larval zebrafish play a similar role, we performed recordings in which 36 distinct optic flow stimuli were presented, followed by the CMN stimulus (Figure S2C). These recordings lasted 20 min. Each optic flow stimulus consisted of either pure translation or rotation, or of a unique combination of TOF and ROF (Figure 3E; also see STAR Methods section prediction of optic flow preference). Using the measured RFs, we applied the matched-filter model to predict the calcium responses to the 36 optic flow stimuli for each neuron. The linear correlation of measured and predicted responses to these distinct optic flow stimuli provides a quantitative measure of the prediction accuracy. The average correlation for the neurons with complex RFs (0.81) was higher than the correlation for the neurons with simple RFs (0.65), which means that the complex RFs on average gave better prediction of the neural responses to optic flow stimuli than the simple RFs (Figures 3F and S2D). We

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

quantified the response specificity to translation versus rotation (TR index; STAR Methods) and found that bimodal translation neurons had a higher specificity for translation (average TR index >0.5; Figure 3G) than rotation neurons for rotation (TR index ~ −0.3). The index distribution for simple RFs did not show strong preference to either translation or rotation stimuli as they were also active during presentations of combined rotation/translation OFs.

Together, these results suggest that the identified complex RFs act as matched filters to process optic flow. Most complex RFs prefer either translation or rotation but not a particular combination of the two. Translation neurons were more abundant than rotation neurons and their responses to combinations of optic flow were more predictable with the matched-filter model (Figure S2D), suggesting that additional unknown stimulus features modulate the responses of rotation neurons. In the following, we will mainly focus on the population properties of the translation RFs.

A topographic map of global translation direction in the pretectum

The encoding of global translation direction is essential for any visually guided locomotion. In the sections above, we showed the encoding algorithm (matched filter) implemented by the complex RFs of optic-flow-sensitive neurons. This allows us to directly infer the preferred translation direction of these neurons as the focus of contraction (FoC) of its best-matching TOF (Figure 4A). Consistent with Figure 2D, most of these neurons prefer translation directions close to the horizontal plane, and another, smaller proportion of neurons prefer vertical TOFs (Figure 4A).

The majority of these neurons were anatomically arranged in the anterior medial cluster (AMC) region[16] of the APT (Figures 4B and S3). These neurons surprisingly formed a topographic map running from left to right across the midline according to preferred translation direction, with medial neurons preferring forward translation. This topographic map is structured according to the ego-motion relative to the surround, i.e., a cursotopic arrangement (lat. "cursus" means "running, direction"). Linear regression showed that the preferred translation direction is strongly correlated with anatomical position ($R^2 = 0.73$; Figure S4F) and that topography is weaker but still present, when removing the effects of left-right (LR) hemisphere organization (correlation = 0.29; Figure S4G). To visualize this cursotopy further, we recorded the neural responses in the AMC to a TOF stimulus, whose direction changed CW at a constant rate. As expected, the active region in the AMC gradually shifted from left to right, when the translation direction was smoothly changing (Video S1).

There are different ways in which a translation-sensitive neuron could detect the direction of TOF. For example, its complex RF could sample the vicinity of the FoC and/or focus of expansion (FoE), or the regions away from FoC and FoE, which contain stronger and more coherent local motion. To identify how the complex translation-sensitive RFs solve this task, we calculated the angular distance between RF mode centers(s) and the preferred translation directions, including both unimodal and bimodal RFs in the analysis. We found that the mode centers of most bimodal translation RFs were located 60°–90° away from the preferred translation direction (Figure 4D), corresponding to the regions with highest speed and strongest local motion coherence in the optic flow field. Furthermore, we found the mode centers of most bimodal translation RFs, which distributed asymmetrically in relation to the body axis (cf. Figure 2C), to distribute symmetrically in relation to the preferred translation directions (Figure 4E). Unimodal complex neurons, however, were distributed more widely and located slightly closer to the FoC or FoE (~51° distance on average).

The responses of translation-sensitive neurons in the AMC were modulated not only by translation direction but also by speed. Neurons were jointly tuned to translation direction and speed (Figure S4) and preferred speeds from approximately 15°/s to 40°/s. To estimate the fidelity of the joint speed and direction representation in the AMC neurons, we computed the discrimination index (DI) for translation direction and speed for each neuron. A higher DI means the neural responses are more discriminable, which in turn, indicates the neuron is more strongly tuned to the variable. We found that almost all neurons were tuned to both translation direction and speed, while the tuning discriminability for translation direction was slightly higher than that for translation speed in most neurons (Figure 4F).

Robust encoding of translation direction

As shown above, neuronal responses to optic flow can be faithfully predicted by the matched-filter algorithm based on the measured pretectal RFs (Figure 3). An essential question to ask is whether the information encoded by neurons with translation RFs is sufficient to estimate the translation direction when distractors or noise are present? In other words, can an animal robustly estimate its ego-motion based on the neural responses of these translation-sensitive neurons in the AMC region?

To answer this question, we first constructed a population decoder to quantitatively determine whether the encoding of translation direction is robust against rotational self-motion interference (STAR Methods section population decoder for self-motion estimation). TOFs are greatly affected by additional ROF components. For example, added rotation will cause an apparent shift of the FoC point and changes of spatial symmetry. This decoder was a variant of the maximum-a-posteriori decoder that used the measured RFs (i.e., using data from Figure 3E) to predict responses to different TOFs. The decoder then estimated the translation direction of the presented optic flow combination by minimizing the squared difference between the recorded responses of the translation-sensitive neurons and the prediction by their RFs (Figure 5A; see STAR Methods for details).

To our surprise, the estimation error for the translation direction was similar with or without ROF interference, which suggests a robust encoding of translation direction (Figure 5A, bottom). Indeed, AMC responses to the pure TOF and the mixed optic flow looked very similar (Videos S1 and S2). In other words, the translation information can be extracted from the mixed optic flow field solely with the information encoded by the translation RFs in the AMC region.

To compare the encoding quality of the speed and direction information, we constructed a separate population decoder (Figures S5A and S5B; see STAR Methods for details). In contrast to the low decoding error for translation direction

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

CellPress

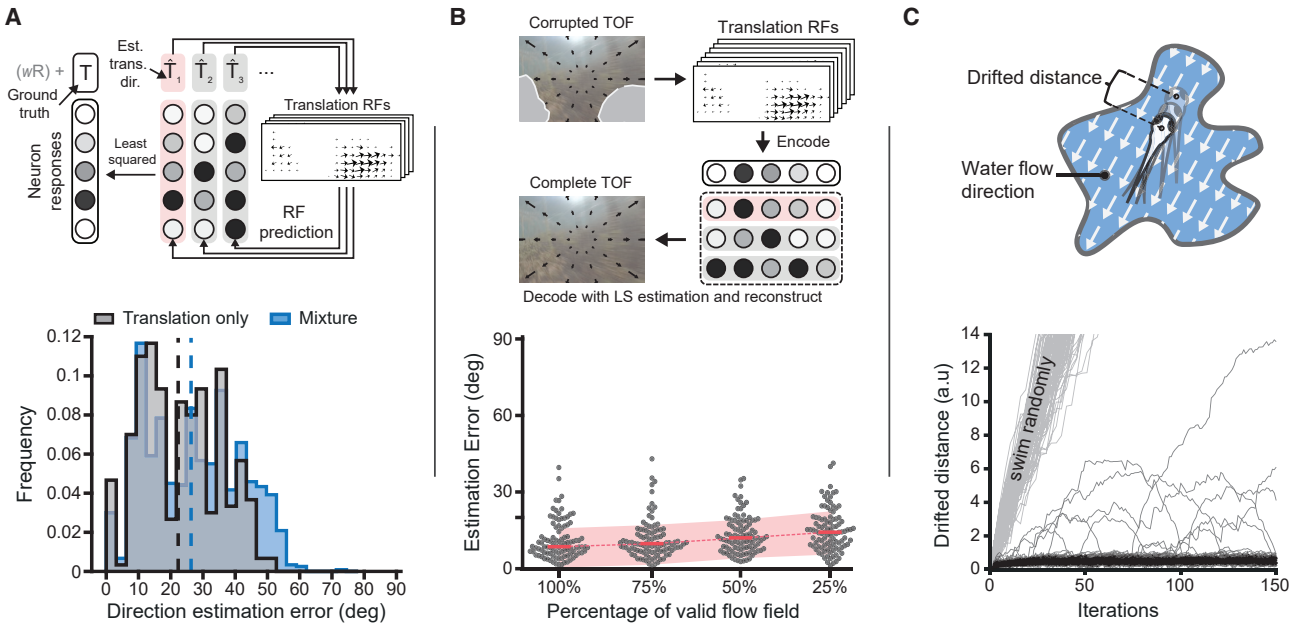


Figure 5. Robust encoding of translational optic flow direction by translation RFs

(A) Top: illustration of the constructed population decoder estimating the stimulus parameters from the responses of optic-flow-sensitive neurons (see also STAR Methods and Figure S5). The "w" and "R" in the top left indicate the rotation weight in relation to translation and rotation parameters, respectively. Est. trans. dir., estimated translation direction. Bottom: the decoder estimated the translation direction for translational optic flow with ("mixture") and without ("translation only") added rotational components. A histogram of estimation errors is shown as well as the median errors (dashed lines). For each data point, a new decoder was built that used 50 randomly selected RFs (out of the 138 RFs; STAR Methods).

(B) Top: illustration of the autoencoder constructed with the translation RFs for determining the direction of corrupted translational optic flow fields (TOFs; STAR Methods). Bottom: bee swarm plots showing the estimation errors for the TOFs with different degrees of visual field corruption. The dashed line and shaded area indicate the median ± SD.

See also Figure S5.

(C) The drifted distance over time of 100 simulated fish (black lines) using the population decoder in (B) to estimate and swim against the direction of ego motion (illustrated in the upper drawing). The gray lines correspond to the drifted distance when the simulated fish swam randomly.

See also Video S3.

(median error: 6.6°), the decoding error for translation speed was high (8.56°/s or 32.6% of the stimulus speed) when compared with the range of encoded speeds.

In the natural environment, optic flow is often corrupted by occlusion, the lack of visual cues or by interfering object motion. Since it is essential for animals to determine their moving direction from corrupted optic flow, we determined to what extent this is possible using translation RFs. For this purpose, we constructed an autoencoder whose encoder and decoder employ the same response functions determined by the corresponding translation RFs (Figure 5B, top).

We first tested the encoding quality for TOFs with 100 different translation directions at 4 completeness levels (25%, 50%, 75%, and 100%), corresponding to the fractional area of the visual field in which unblocked optic flow is available. For example, 25% indicated the optic flow in 25% of the visual field is visible. For each TOF, an autoencoder was constructed with 50 RFs randomly sampled from the total 329 translation RFs (see STAR Methods for details). As expected, the estimation error gradually increased with the decreasing completeness levels; however, even for the 25% completeness level in which most of the optic flow field is missing, the median estimation error was still below 20°. It was only 5.7° higher than the one for the 100% completeness level (8.7°; Figure 5B, bottom), suggesting that the corresponding RFs are well prepared to encode self-motion direction in environments with sparse motion information. We also tested the encoding robustness against moving objects, which corrupt optic flow fields with wrong motion information. For this type of corruption, the decoding was less robust, especially when half or more than half of the visual field was covered by moving objects (Figure S5E). This poor performance was not surprising, since it is hard to distinguish the cues for global motion and object motion purely from the optic flow fields without additional information.

The robust encoding of translation direction is critical for many locomotor behaviors, e.g., for self-stabilizing behavior in water currents.[28] The continuous change of the water current direction and the relative distance to the surround alter the optic flow continuously, which requires fish to constantly adjust their swimming direction to maintain their body position. We were curious to see, whether these translation RFs are capable of such a complicated task. We therefore constructed 100 randomly shaped "ponds." In each of these simulated ponds, the water flowed at constant speed (1 a.u./iteration) in a random direction. The simulated fish then swam at the same speed into the directions estimated by the autoencoder from the instantaneous optic

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

flow fields (Figure 5C; also see example fish in Video S3). In most simulations, the total drifted distances were smaller than 1 a.u., which suggested that the translation RFs accurately encoded sufficient information in this highly dynamic process to determine appropriate translation directions for self-stabilization behavior.

Distinct optokinetic and OMRs to translational and rotational components in complex optic flow fields

Earlier, we asserted that without systematic characterization of neural mechanisms, we cannot even answer how fish will respond to the optic flow during a bank turn as it consists of both translation and rotation (Figure 6A). The revealed matched-filter algorithm provides insight in the self-motion extraction from optic flow in zebrafish larvae. The robust encoding of translation direction (Figure 5A; Video S2) against the interference of ROF component strongly suggests translation and rotation information from the optic flow input are segregated in the zebrafish brain. It is thus very likely that fish will respond separately and simultaneously to the translation and rotation related optic flow components during any movements involving both translation and rotation.

To test this, we selected the OMRs and OKRs as two behavior readouts as they are tightly linked to optic flow perception and represent prominent stabilization behaviors in zebrafish.[16,28] We recorded the OKR and OMR of 24 freely swimming fish to pure forward/backward translational and CW/counterclockwise (CCW) ROF, as well as to combinations of both (Figures 6B and 6C).

If fish failed to decompose translational and rotational components, we would expect inappropriate behavioral responses, for example, the occurrence of rotational behavior to TOF. In our experiments, fish responded to pure rotation with both OMR and OKR, as indicated by their circular swimming trajectories and conjugate eye movements (Figures 6C and S6). For the pure translation optic flow, fish swam with the optic flow direction, while clear conjugate or vergence eye movements were absent.

If fish were unable to extract both translational and rotational self-information from the combinatory optic flow, their responses should be similar to the responses evoked by pure translational or ROF. Interestingly, when combinations of translation and rotation stimuli were presented, fish clearly swam with the translation direction while performing conjugate eye movements in the rotation direction, indicating the perception of both translation and rotation components in the optic flow stimuli (Figures 6C and S6).

Since OKR during translation stimuli may be disrupted by changes of head direction in freely swimming fish, we repeated the experiment in semi-restrained fish, which could not swim but were able to move their eyes (Figure 6D). In contrast to the results in freely swimming fish, the semi-restrained fish performed clear vergence eye movements to pure TOFs. For optic flow combinations, both semi-restrained and freely swimming moving fish showed conjugate eye movements without significant vergence components (Figure 6E). Together, this behavioral evidence suggest that translation and rotation components are perceived separately and simultaneously by zebrafish.

There are at least two local features that are different between the pure translational and the mixed optic flow: the FoC/FoE position and the average motion speed of the left and right visual hemi-fields (cf. upper right illustration in Figure 6E). The head direction invariant OKR implied that the distinct processing of translation optic flow with or without rotational components does not rely on particular local motion features. To test this, we constructed two variants of TOF. In the ''slanted translation'' variant, the translation direction of the pure TOF was shifted, so that its FoC was at the same position as the one in the mixture of translation- and rotation-induced optic flow. The average speed in the left and right halves of the ''asymmetric translation'' optic flow fields was altered to mimic the different speed in the mixture optic flow field. The OKRs to these stimuli were slightly different from the ones to pure translation optic flow; however, none of them could be classified as conjugate eye movements (Figure 6E).

We quantified the extent of conjugate versus vergence eye movements using a vergence index (Figure 6F, 1 = pure vergence movement; 0 = pure conjugate movement; see STAR Methods for details). The vergence index for the translation optic flow combined with a rotation component was significantly shifted toward conjugate eye movements in comparison with all other translation stimuli, while the vergence indices for the various translation stimuli were similar to each other (Figure 6F; Wilcoxon rank-sum tests). Furthermore, each vergence index (except for the ''asymmetric translation'' stimulus) was slightly shifted toward conjugacy in comparison with what would be expected based on the distribution of local velocity vectors in the optic flow field (see black crosses in Figure 6F, which show the bootstrapped null distribution). Together, these results show that the animal cannot be tricked into processing rotational motion when only certain local aspects of rotational flow are implemented in overall translational flow patterns.

DISCUSSION

Via the systematic characterization of the fine-grained RF structure of motion-sensitive neurons in the OT and pretectal area, we revealed how self-motion information can be extracted from the complex RFs, and how the encoding properties may be explained by the matched-filter model. We furthermore demonstrated the robust encoding of self-translation direction and speed in the translation-sensitive neurons in the pretectal AMC region. Finally, our behavioral results show that zebrafish optokinetic and OMRs draw on the decomposition of rotational and translational information.

Complex RFs implement the optic flow detection algorithm

To our knowledge, this is the first study that systematically estimated the detailed RF structure of optic-flow-processing neurons in a vertebrate animal model. The feasibility to estimate the 3,926 motion RFs reported here, hinged on the availability of a fast, multiplexed RF estimation technique developed for calcium imaging.[26] The RF structures of the global motion-sensitive neurons revealed in this study showed remarkable similarity to translational and ROFs (Figure 3C). This observation is consistent with the ''matched-filter'' model proposed in flies, which suggests optic-flow-sensitive neurons extract self-motion information by matching their RFs with impinging optic flow fields.[15,23,29] The calcium responses of our complex RF neurons,

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

CellPress

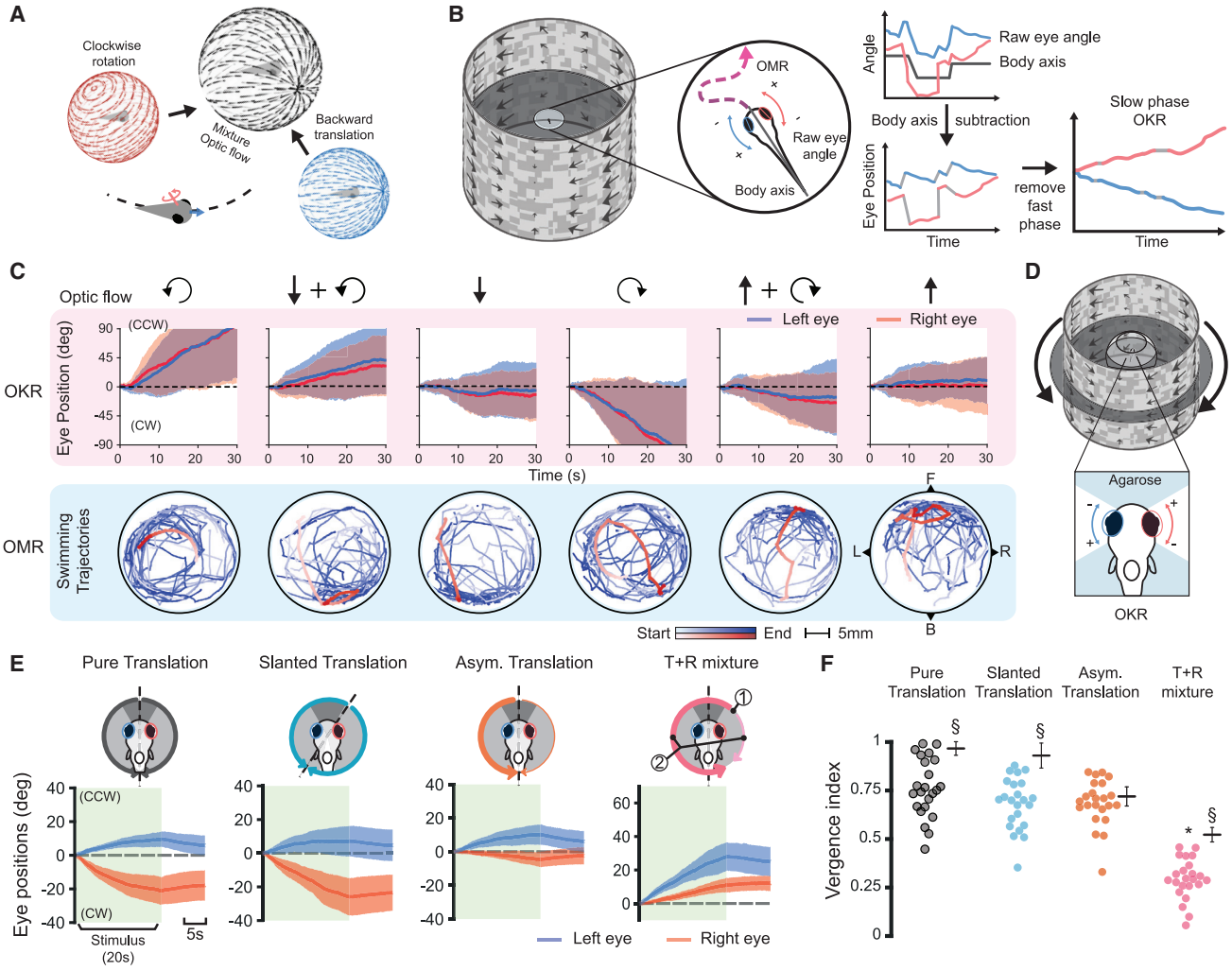


Figure 6. Zebrafish decompose mixed rotation and translation optic flow in optokinetic and optomotor responses

(A) Illustration of the mixture optic flow field resulting from the counterclockwise (CCW) turning of a forward moving fish.

(B) Setup for simultaneous OKR (eye movements) and OMR (swimming trajectory) recording for freely swimming fish (see also STAR Methods and Figure S6). For OKR (right), saccades were removed to analyze cumulative slow phase eye movements.

(C) The OKR and OMR upon optic flow stimulation containing translation and/or rotation components in 24 fish (black arrows on top indicate the OF stimulus composition; solid line, mean; shaded area, SD across all fish). One example swimming trajectory is colored in red for each stimulus. "F," "B," "L," and "R": the forward, backward, left, and right direction, respectively.

(D) OKR recording setup for agarose-mounted "semi-restrained" fish in a spherical water container. The agarose around the eyes was removed to allow free eye movement.

(E) The eye position traces (mean ± SD) across 18 trials recorded in five semi-restrained fish. The OF stimuli are illustrated on top. Dark gray corresponds to binocular overlap region, and the curvy arrows in (D) and (E) indicate the average motion in each hemi visual field. Arrow thickness corresponds to the relative stimulus speed. ① and ② label two local features present when adding rotation on top of translation OF: the FoE shifts and speed in the two visual hemi-fields is unequal.

(F) Ocular vergence index for different OF patterns (based on eye traces in E). For comparison, we also simulated the vergence index (black mean ± SD bars) expected for independently moving eyes responding to the average stimulus directions and speeds in each hemi-field (using bootstrapping of local stimulus statistics; STAR Methods). 1 and 0 correspond to pure convergent/divergent and pure conjugate movement, respectively. The asterisk indicates significantly lower vergence indices compared with the other three groups. The "§" indicates a significantly lower measured ocular vergence index in comparison with the corresponding local flow statistic. The rank-sum test ($\alpha = 0.05$) and Bonferroni correction were used to assess significant differences.

especially the ones with translation RFs, are well described by this model (Figures 3F and S2D), and the level of remaining unexplained response modulation is small for most of the complex RF neurons. Most complex RFs are bimodal RFs (Figure 3B) and it is possible that the bimodal RFs of these neurons are inherited from pairs of presynaptic unimodal RF neurons,[30] which may be confirmed by systematic connectivity characterization in the future.

In contrast, the responses of the simple RF neurons frequently differed from the predictions of the matched-filter model

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

(Figures 3F and S2D), and the precise choice of thresholds in our analysis did not appear to have a major effect on the matched-filter properties of neurons (data not shown). One possibility is that many of these "simple" neurons had non-linear RF structure (i.e., suppressive components) remaining undetected in our RF estimation. Indeed, previous studies reported that about 80% of the local motion-sensitive neurons in zebrafish OT and salamander retinal ganglion cells (RGCs) were suppressed by wide-field motion.[19,31] An alternative (and not mutually exclusive) explanation for the observed prediction discrepancy is that our simple RF neurons carry out functions in the visual system, which are only distantly related to optic flow encoding. Further investigation is needed to characterize inhibitory components and functions of simple RF neurons in the brain.

A topographic map of preferred translation directions in the AMC pretectal region

Within the pretectal AMC region, translation-selective neurons are roughly arranged according to preferred translation directions in a single topographic map that spans both pretectal hemispheres (Figures 4B, S4F, and S4G; Videos S1 and S2). Previous reports already revealed differential directional processing across pretectal locations and related these direction-selective neurons to the triggering of OMRs.[16–19,32]

These results agree with our findings despite major differences in employed visual stimulus setups. In these studies,[17,32] motion stimuli were projected on a flat screen from below, which can be interpreted as the bottom part of a TOF. For example, the nasal-temporal-ward motion covering the bottom visual field can be interpreted as the bottom part of a TOF resulting from forward self-motion. Using this interpretation, the topographic map of translation direction found in this study (Figure 4B) is consistent with the direction map described in the previous study[17] (Ibid. Figure 3A) in terms of the PD in different regions of the AMC. Furthermore, our results show that this pretectal structure encodes the translation axes of optic flow rather than local motion directions for each eye. Unfortunately, RF structures in the peripheral lower visual field ($<-40^{\circ}$ elevation) could not be estimated due to the hardware limitation in our study. A whole-field stimulus display setup will be desired for estimating the complete RF structure in the entire visual field.

Besides translation direction, speed is another critical factor for locomotor behavior. Larval zebrafish can adjust their swimming behavior to match the speed of wide-field visual motion stimuli.[33,34] Here, we show that most translation-sensitive neurons in the AMC region are jointly tuned to both translation direction and speed (Figure 4F). The speed information in TOF is usually conjugated with the depth and location of stimuli in the visual field.[21] Whether and how translation-sensitive neurons in the AMC regions untangle this information for determining self-motion speed requires further investigation.

Represented optic flow axes

We noticed some invariant properties across RFs in the OT and pretectum: (1) the PDs for both the unimodal and bimodal RFs are biased toward either horizontal or vertical motion, with only a few neurons responding to oblique motion in Figures 2D and S2A;[18,35,36] (2) bimodal vertical neurons together encode translation and rotation, whereas most bimodal horizontal neurons encode translation (Figure 2D); (3) the modes of bimodal translation RFs are mostly distributed symmetrically with respect to the preferred translation direction instead of the body axis (Figures 2C and 4E); and (4) most translation RFs sample regions to the side of the preferred translation direction, where motion coherence is high (Figure 4D).

It is unclear at this point what the specific advantages of this layout are. The layout results in a relatively even distribution of neuronal PDs in the horizontal plane, while translation neurons preferring oblique translation directions outside the horizontal plane or the vertical axis are absent. One possibility is that this representation serves to improve the coding efficiency by balancing the needed number of active neurons (sparseness) and the needed total neuron number: the orthogonal representations could form a mathematical basis, so that oblique directions could be encoded by the activity of (at least) two neurons with orthogonal PDs. Our decoder simulations show that the relatively rich representation of horizontal directions can be beneficial for targeted locomotor responses in the horizontal plane, e.g., during OMR (Figure 5C). However, spontaneous swimming as well as prey capture behavior includes frequent locomotor deviations from the horizontal plane[37,38] as well. The surprising scarcity of representations of yaw rotation suggests that it is negligible in visual brain areas of the zebrafish. Possibly, inhibitory outputs of translation-selective neurons or the activity in the vestibular system could be integrated in downstream neurons to signal the presence of yaw rotation. Furthermore, each eye can, to some extent, correct for horizontal image slip by itself, potentially making a binocular rotation representation dispensable.

Decomposition of translation and rotation self-motion components

While evidence for separate processing of translation and rotation information in the corresponding neural correlate has been presented before,[12,16,18,39] our study shows that translation/rotation-sensitive neurons are not exclusively responsive to pure TOF/ROF but also respond to the mixed optic flows in a predictable manner (Figure 3). Furthermore, the population encoding of translation direction is robust against rotation interference (Figure 5A; Video S2), which indicates that the zebrafish pretectum separates translational and rotational self-motion components and is biased toward the identification of translation components. In contrast to translation RFs, the responses of rotation-sensitive neurons to mixed optic flow stimuli are strongly modulated by the non-preferred (translation) components and poorly explained by their RF structure (Figure S2D). It remains to be answered whether these rotation-sensitive neurons in zebrafish are similar to the neurons jointly representing translation and rotation components in primate parietal lobes[12] and what role they do play in optic flow processing.

The decomposition of optic flow requires the translational and rotational self-motion information not only to be separated but also to be utilized differentially in behavior. Our behavioral experiments revealed evidence of this decomposition, since fish swam toward the corresponding translation direction implemented in the optic flow field (OMR; Figure 6C) while spinning both eyes in the direction of rotation (OKR; Figures 6C, 6E, and 6F). The OKR is more sensitive to the rotation component of a combined

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

optic flow stimulus (Figure 6F), suggesting that OKR and OMR might show performance differences for different types of optic flow. Our experiments on isolated local features (Figures 6E and 6F) suggest that the integration of local features but not a specific local feature is essential for the optic flow processing in zebrafish.

We hypothesize that our identified population of pretectal neurons form a visual encoding pivot, where complex RFs directly extract and convey the translation direction to the premotor areas. It will be interesting to determine the full neural circuit for this optic flow decomposition in the future, including the connectivity of translation-sensitive neurons in AMC. [30]

In summary, we have systematically characterized the RF properties in zebrafish visual brain areas that help us to reveal a neural mechanism for optic flow processing. Our results suggest that individual neurons form matched filters to extract translation and rotation self-motion components from optic flow. We identified an anatomically ordered arrangement of different functional cell types, which, as a population, decompose rotational and translational self-motion components of the animal. Future work is needed to show, how these cellular building blocks are wired together in a hierarchical circuit to mediate visually driven reflexes and voluntary behaviors.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials Availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
- METHOD DETAILS
  - Visual stimulation setup
  - Contiguous motion noise stimulus
  - Optic flow stimuli design
  - *In vivo* two-photon calcium imaging
  - Recording behavior in freely moving animals
  - Eye movement recording in mounted fish
  - Stimulus design for behavior experiments
- RETROGRADE LABELLING OF NMLF
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Calcium data pre-processing
  - High throughput motion RF estimation
  - Fitting optic flow fields to RFs
  - Prediction of optic flow preference
  - Translation direction and speed tuning
  - Population decoder for self-motion estimation
  - Assessing vergence of optokinetic responses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.04.048.

ACKNOWLEDGMENTS

We thank Thomas Nieß (glassblower shop, University of Tübingen) and Klaus Vollmer (fine mechanics workshop, University of Tübingen) for technical support, and Herwig Baier (MPI of Neurobiology) for providing the Tg(elavl3:nls-GCaMP6s)mpn400 zebrafish line. We thank Holger Krapp (Imperial College London), Eva Naumann (Duke School of Medicine), Julian Hinz (Friedrich Miescher Institute for Biomedical Research), and members of the laboratory for feedback on the manuscript. This work was funded by the Deutsche Forschungsgemeinschaft (DFG) grants EXC307 (CIN – Werner Reichardt Centre for Integrative Neuroscience) and INST 37/967-1 FUGG, and Human Frontier Science Program (HFSP) Young Investigator grant RGY0079.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.Z. and A.B.A.; methodology, Y.Z., R.H., and A.B.A.; software, Y.Z. and R.H.; formal analysis, Y.Z. and A.B.A.; investigation, Y.Z., R.H., and W.N.; writing – original draft, Y.Z. and A.B.A.; writing – review & editing, Y.Z., R.H., W.N., and A.B.A.; visualization, Y.Z.; supervision, A.B.A. and Y.Z.; funding acquisition, A.B.A.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 4, 2022
Revised: March 14, 2022
Accepted: April 14, 2022
Published: May 11, 2022

REFERENCES

1. Gibson, J.J. (1954). The visual perception of objective motion and subjective movement. Psychol. Rev. *61*, 304–314.
2. Lappe, M., Pekel, M., and Hoffmann, K.-P. (1998). Optokinetic eye movements elicited by radial optic flow in the macaque monkey. J. Neurophysiol. *79*, 1461–1480.
3. Nakayama, K., and Loomis, J.M. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: a hypothesis. Perception *3*, 63–80.
4. Britten, K.H. (2008). Mechanisms of self-motion perception. Annu. Rev. Neurosci. *31*, 389–410.
5. Festl, F., Recktenwald, F., Yuan, C., and Mallot, H.A. (2012). Detection of linear ego-acceleration from optic flow. J. Vision *12*, 10.
6. Zosh, W.D., Duchon, A.P., Sahuc, S., Duchon, A.P., and Sahuc, S. (2001). Optic flow is used to control human walking. Nat. Neurosci. *4*, 213–216.
7. Srinivasan, M., Zhang, S., Lehrer, M., and Collett, T. (1996). Honeybee navigation en route to the goal: visual flight control and odometry. J. Exp. Biol. *199*, 237–244.
8. Warren, W.H., and Hannon, D.J. (1988). Direction of self-motion is perceived from optical flow. Nature *336*, 162–163.
9. Duffy, C.J., and Wurtz, R.H. (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J. Neurophysiol. *65*, 1329–1345.
10. Duffy, C.J., and Wurtz, R.H. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. J. Neurosci. *15*, 5192–5208.
11. Luo, J., He, K., Andolina, I.M., Li, X., Yin, J., Chen, Z., Gu, Y., and Wang, W. (2019). Going with the flow: the neural mechanisms underlying illusions of complex-flow motion. J. Neurosci. *39*, 2664–2685.
12. Sunkara, A., DeAngelis, G.C., and Angelaki, D.E. (2016). Joint representation of translational and rotational components of optic flow in parietal cortex. Proc. Natl. Acad. Sci. USA *113*, 5077–5082.
13. Tanaka, K., and Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal

part of the medial superior temporal area of the macaque monkey. J. Neurophysiol. *62*, 626–641.

14. Wylie, D.R.W. (2000). Binocular neurons in the nucleus lentiformis mesencephali in pigeons: responses to translational and rotational optic flowfields. Neurosci. Lett. *291*, 9–12.
15. Krapp, H.G., and Hengstenberg, R. (1996). Estimation of self-motion by optic flow processing in single visual interneurons. Nature *384*, 463–466.
16. Kubo, F., Hablitzel, B., Dal Maschio, M., Driever, W., Baier, H., and Arrenberg, A.B. (2014). Functional architecture of an optic flow-responsive area that drives horizontal eye movements in zebrafish. Neuron *81*, 1344–1359.
17. Naumann, E.A., Fitzgerald, J.E., Dunn, T.W., Rihel, J., Sompolinsky, H., and Engert, F. (2016). From whole-brain data to functional circuit models: the zebrafish optomotor response. Cell *167*, 947–960.e20.
18. Wang, K., Hinz, J., Haikala, V., Reiff, D.F., and Arrenberg, A.B. (2019). Selective processing of all rotational and translational optic flow directions in the zebrafish pretectum and tectum. BMC Biol. *17*, 29.
19. Wang, K., Hinz, J., Zhang, Y., Thiele, T.R., and Arrenberg, A.B. (2020). Parallel channels for motion feature extraction in the pretectum and tectum of larval zebrafish. Cell Rep. *30*, 442–453.e6.
20. Dahmen, H.-J., Franz, M.O., and Krapp, H.G. (2001). Extracting egomotion from optic flow: limits of accuracy and neural matched filters. In Motion Vision: Computational, Neural and Ecological Constraints, J.M. Zanker, and J. Zeil, eds. (Springer), pp. 143–168.
21. Koenderink, J.J., and van Doorn, A.J. (1987). Facts on optic flow. Biol. Cybern. *56*, 247–254.
22. Longuet-Higgins, H.C., and Prazdny, K. (1980). The interpretation of a moving retinal image. Proc. R. Soc. Lond. B Biol. Sci. *208*, 385–397.
23. Franz, M.O., and Krapp, H.G. (2000). Wide-field, motion-sensitive neurons and matched filters for optic flow fields. Biol. Cybern. *83*, 185–197.
24. Huston, S.J., and Krapp, H.G. (2008). Visuomotor transformation in the fly gaze stabilization system. PLoS Biol. *6*, e173.
25. Wertz, A., Haag, J., and Borst, A. (2009). Local and global motion preferences in descending neurons of the fly. J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol. *195*, 1107–1120.
26. Zhang, Y., and Arrenberg, A.B. (2019). High throughput, rapid receptive field estimation for global motion sensitive neurons using a contiguous motion noise stimulus. J. Neurosci. Methods *326*, 108366.
27. Bianco, I.H., Kampff, A.R., and Engert, F. (2011). Prey capture behavior evoked by simple visual stimuli in larval zebrafish. Front. Syst. Neurosci. *5*, 101.
28. Roeser, T., and Baier, H. (2003). Visuomotor behaviors in larval zebrafish after GFP-guided laser ablation of the optic tectum. J. Neurosci. *23*, 3726–3734.
29. Krapp, H.G., Hengstenberg, B., and Hengstenberg, R. (1998). Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly. J. Neurophysiol. *79*, 1902–1917.
30. Kramer, A., Wu, Y., Baier, H., and Kubo, F. (2019). Neuronal architecture of a visual center that processes optic flow. Neuron *103*, 118–132.e7.
31. Kühn, N.K., and Gollisch, T. (2016). Joint encoding of object motion and motion direction in the salamander retina. J. Neurosci. *36*, 12203–12216.
32. Yildizoglu, T., Riegler, C., Fitzgerald, J.E., and Portugues, R. (2020). A neural representation of naturalistic motion-guided behavior in the zebrafish brain. Curr. Biol. *30*, 2321–2333.e6.
33. Portugues, R., Haesemeyer, M., Blum, M.L., and Engert, F. (2015). Whole-field visual motion drives swimming in larval zebrafish via a stochastic process. J. Exp. Biol. *218*, 1433–1443.
34. Severi, K.E., Portugues, R., Marques, J.C., O'Malley, D.M., Orger, M.B., and Engert, F. (2014). Neural control and modulation of swimming speed in the larval zebrafish. Neuron *83*, 692–707.
35. Hunter, P.R., Lowe, A.S., Thompson, I.D., and Meyer, M.P. (2013). Emergent properties of the optic tectum revealed by population analysis of direction and orientation selectivity. J. Neurosci. *33*, 13940–13945.
36. Lowe, A.S., Nikolaou, N., Hunter, P.R., Thompson, I.D., and Meyer, M.P. (2013). A systems-based dissection of retinal inputs to the zebrafish tectum reveals different rules for different functional classes during development. J. Neurosci. *33*, 13946–13956.
37. Bolton, A.D., Haesemeyer, M., Jordi, J., Schaechtle, U., Saad, F.A., Mansinghka, V.K., Tenenbaum, J.B., and Engert, F. (2019). Elements of a stochastic 3D prediction engine in larval zebrafish prey capture. eLife *8*, e51975.
38. Ehrlich, D.E., and Schoppik, D. (2017). Control of movement initiation underlies the development of balance. Curr. Biol. *27*, 334–344.
39. Wylie, D.R. (2013). Processing of visual signals related to self-motion in the cerebellum of pigeons. Front. Behav. Neurosci. *7*, 4.
40. Kunst, M., Laurell, E., Mokayes, N., Kramer, A., Kubo, F., Fernandes, A.M., Förster, D., Dal Maschio, M., and Baier, H. (2019). A cellular-resolution atlas of the larval zebrafish brain. Neuron *103*, 21–38.e5.
41. Förster, D., Dal Maschio, M., Laurell, E., and Baier, H. (2017). An optogenetic toolbox for unbiased discovery of functionally connected cells in neural circuits. Nat. Commun. *8*, 116.
42. Lister, J.A., Robertson, C.P., Lepage, T., Johnson, S.L., and Raible, D.W. (1999). nacre encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. Development *126*, 3757–3767.
43. dal Maschio, M., Donovan, J.C., Helmbrecht, T.O., and Baier, H. (2017). Linking neurons to network function and behavior by two-photon holographic optogenetics and volumetric imaging. Neuron *94*, 774–789.e5.
44. 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.
45. Reiser, M.B., and Dickinson, M.H. (2008). A modular display system for insect behavioral neuroscience. J. Neurosci. Methods *167*, 127–139.
46. Rinner, O., Rick, J.M., and Neuhauss, S.C.F. (2005). Contrast sensitivity, spatial and temporal tuning of the larval zebrafish optokinetic response. Invest. Ophthalmol. Vis. Sci. *46*, 137–142.
47. Euler, T., Hausselt, S.E., Margolis, D.J., Breuninger, T., Castell, X., Detwiler, P.B., et al. (2009). Eyecup scope—optical recordings of light stimulus-evoked fluorescence signals in the retina. Pflugers Arch. *457*, 1393–1414.
48. Wang, K., Arrenberg, B., Hinz, J., and Arrenberg, A.B. (2021). Reduction of visual stimulus artifacts using a spherical tank for small, aquatic animals. Sci. Rep. *11*, 3204.
49. Thiele, T.R., Donovan, J.C., and Baier, H. (2014). Descending control of swim posture by a midbrain nucleus in zebrafish. Neuron *83*, 679–691.
50. DeAngelis, G.C., and Uka, T. (2003). Coding of horizontal disparity and velocity by MT Neurons in the alert macaque. J. Neurophysiol. *89*, 1094–1111.
51. Abbott, L.F., and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. Neural Comput. *11*, 91–101.

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
10000 MW Texas Red dextran	Thermo Fisher	https://www.thermofisher.com/
Deposited data		
Max Planck Zebrafish Brain Atlas	Kunst et al.[40]	https://fishatlas.neuro.mpg.de
Processed data for generating the figures contained in this manuscript	This paper; G-Node GIN	https://doi.gin.g-node.org/10.12751/g-node.hkzu3l
Experimental models: Organisms/strains		
Tg(elavl3:nls-GCaMP6s)mpn400	Förster et al.[41]	ZFIN ID: ZDB-ALT-170731-37
mitfa	Lister et al.[42]	ZFIN ID: ZDB-GENE-990910-11
Software and algorithms		
MATLAB (data analysis)	MathWorks	https://www.mathworks.com/products/matlab.html
MOM Computer System (MCS)	Sutter instruments	https://www.sutter.com/MICROSCOPES/mcs.html
Code for generating the figures contained in this manuscript	This paper; G-Node GIN	https://doi.gin.g-node.org/10.12751/g-node.hkzu3l

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Aristides B. Arrenberg (aristides.arrenberg@uni-tuebingen.de).

Materials Availability

This study did not generate new unique reagents

Data and code availability

The scripts and the required pre-processed dataset for the figures and supplementary figures included in the current study are available in a public repository (https://doi.org/10.12751/g-node.hkzu3l).

The raw calcium imaging data and animal behavior videos supporting the current study have not been deposited in a public repository because of their large size but are available from the corresponding author on request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

Zebrafish (Danio rerio) were maintained on a 14/10h light–dark cycle at 28°C. The transgenic zebrafish line *Tg(elavl3:nls-GCaMP6s)mpn400* was used for the *in vivo* two-photon calcium imaging.[43] The top 10 – 20% of larval zebrafish with high calcium indicator expression levels and strong optokinetic responses (OKR) to moving gratings were selected from all hatched fish at 4-day post-fertilization (dpf) for *in vivo* calcium imaging experiments.

5-7 dpf zebrafish larvae carrying mutations in the *mitfa* gene (nacre) were used for OKR and optomotor responses (OMR) recordings. Again, the top 10-20% larval zebrafish sensitive to visual motion stimuli indicated by their strong OKR to moving gratings were selected at 4 dpf for these behavioral experiments.

All behavioral and *in vivo* calcium imaging experiments were performed on 5-6 dpf larval zebrafish at room temperature.

All animal experiments were licensed by the local authorities (Regierungspräsidium Tübingen) in accordance with German federal law and Baden-Württemberg state law.

METHOD DETAILS

Visual stimulation setup

All visual stimuli used in this study were displayed on a custom-built cylindrical green LED arena (160 mm height, inner diameter = 198 mm) covering 336°-by-80° of the visual field (-168° to 168° in azimuth; -40° to 40° in elevation). The 2 halves of LED arena

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

consisted of 224 square LED tiles (8 rows × 14 columns × 2 arena halves); each tile had 8×8 evenly distributed LEDs emitting at 570nm (Kingbright TA08-81CGKWA). Each LED covered ~1.5 visual degrees on average, note the coverage on vertical direction varied slightly with the latitude of LEDs. All stimuli displayed in the arena had 100% contrast as the LEDs in the stimuli were either on or off. All LED tiles were covered with a diffusion filter foil (LEE no. 252, article 595-1780-2520) and a high-pass filter foil (LEE no. 779, article 595-1700-7790, castinfo.de, Hagen, Germany) for light homogenization and light interference reduction in calcium imaging respectively.

The custom software developed in previous studies[44,45] was used for stimulus uploading and display control. To minimize the interference of LED light with the calcium imaging, the LEDs were flickering in a duty cycle (~1 ms) to be switched on only during the flyback time (approx. 160 μs) of the scanning mirrors in-between line scans.

Contiguous motion noise stimulus

The contiguous motion noise (CMN) stimulus for motion RF estimation was constructed as described in our previous study.[26] Briefly, the optic flow in the CMN stimulus was determined by a three-dimensional (8 rows × 28 columns x 12000 frames) motion vector noise matrix, in which motion vectors were spatiotemporally correlated. The vector length and direction indicate the motion speed and direction in the corresponding local region. The spatiotemporal correlation in the noise was achieved by convolving a white vector noise matrix with a multidimensional Gaussian kernel ($\mu_{x,y,t} = 0$, $\sigma_{x,y} = 18.78$ visual degrees, $\sigma_t = 0.33$ seconds). The spatiotemporal correlation in our stimulus may introduce bias to the RF estimation, i.e. overestimation of the local coherence in the RF. The spatiotemporal extent of this bias was quantified as the contiguous radius (CR) of the CMN stimulus for which the pairwise vector correlation surpassed 0.1. The CR is calculated as:

$$CR = 2\sigma\sqrt{-\ln(0.1)} \approx 3.03\sigma$$

The CR in spatial and temporal domain are therefore 57 (3.03×18.78) visual degrees and 1 (3.03×0.33) second respectively.

The noise vector matrix was then transformed to a 64 (row) × 224 (column) × 12000 (frame) binary movie displaying at 15 frame-per-second (fps) with the Reichardt encoder described previously.[26] Each motion vector in the noise vector matrix was approximated by the displacement of the binary noise pattern displayed on an 8×8 LED patch at the corresponding position on the arena. The spatial frequency of the 224 binary noise patterns (one for each LED patch) was set to 0.1 cycles per visual degree which is optimized for visual motion perception in larval zebrafish.[46]

The speed distribution of the CMN stimulus follows the Rayleigh distribution:

$$s(x) = \frac{x}{\sigma^2} e^{\frac{-x^2}{2\sigma^2}},$$

where $\sigma = 12.35$ visual degrees/second. Therefore, the speed in 95% of the pixels in our CMN stimulus movie was contained in the range between 0 to 30 visual degree/second.

Optic flow stimuli design

Similar to the CMN stimulus, the optic flow stimuli were designed in vector field forms with the same spatial resolution (8×28) as in the CMN stimulus, and they were converted to binary noise movies by the same Reichardt encoder to be displayed on the LED arena. The rotational $(\vec{V}_{rot})$ and translational optic flow vector fields $(\vec{V}_{trans})$ were constructed as following:

$$\vec{V}_{rot} = -\vec{P} \times \vec{R},$$

$$and\ \vec{V}_{trans} = \vec{T} - \left(\vec{P} \cdot \vec{T}\right)\vec{P}$$

$\vec{P}$ is a unit vector point from the observer (the origin) to the location P in the visual field. $\vec{T}$ and $\vec{R}$ are the unit vectors indicating the translation direction and the rotation axis of the translational and the rotational optic flow fields respectively. The operations $< \times >$ and $<\cdot>$ are the cross and dot products of two vectors respectively. And a mixture optic flow containing both translation and rotation components can be obtained as:

$$\vec{V}_{mix} = w_{trans}\vec{V}_{trans} + w_{rot}\vec{V}_{rot},$$

where the w_{trans} and w_{rot} represent the relative strength of the translation and rotation components in the mixed flow field.

The speed of motion cues in a translational optic flow field depends on both their locations in the optic flow field and the distance to the observer. As this study is not focused on the distance effect, all motion cues in our optic flow fields were set to be equally distant from the observer. To correct the projection distortion of the optic flow fields on the cylindrical LED arena, the optic flow fields calculated above $(\vec{V}_{raw})$ were transformed as following:

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

$$\vec{V}_{corrected} = \mathbf{D}\left(\vec{V}_{raw} - \left(\vec{V}_{raw} \cdot \vec{N}\right)\vec{N}\right),$$

$$\text{where } \vec{N} = \frac{(x, y, 0)}{|(x, y, 0)|},$$

$$\text{and } \mathbf{D} = \frac{r}{cos(\theta)} = \frac{r}{\sqrt{1 - z^2}},$$

where the *x, y, z* are the LED locations in the Cartesian coordinate system where the origin is the location of the observer, and θ is the corresponding elevation angle.

In the translational optic flow fields, the motion speed is maximum at the locations S which are perpendicular to the translation direction ($\vec{S} \perp \vec{T}$). For convenience, the translation speed of a translational optic flow field was defined as the speed in these locations (color-shaded circles in the illustrations of Figure S4A).

In vivo two-photon calcium imaging

The selected fish were immobilized in 1.7% low melting-temperature agarose (wt/vol, E3 medium) at the tip of a transparent plastic triangle stage (tip angle < 45%, Figure S2E) on the experiment day. The triangle stage with the mounted fish was later placed in the centers of a spherical glass container filled with transparent E3 medium. Then the water container was fixed on a metal holder placed in the centers of the LED arena. The body position and orientation of the mounted fish was adjusted so it was centered in the LED arena with the dorsal side facing up (roll correction) and the nose pointing straight to the front (0° azimuth and elevation) without any tilting.

The calcium activities in neuron somata were recorded with a movable objective microscope two-photon setup (Sutter Instruments, Novato, California, USA)[47] coupled to a Coherent Vision-S Ti-Sa laser. The emitted fluorescence light of the calcium indicator protein, GCaMP6s, triggered by the excitation laser beam at 920 nm (prepulse compensation: 9756 fs^2, laser power = ~33 mW at object plane) was collected with a 20x/1.0 Zeiss objective using an emission dichroic with bandpass 500–550 nm. The image time series were recorded using MScan software (Sutter instruments). The calcium images acquisition rate was 2 fps at 512 × 512 $pixel^2$ and a magnification of 2 (0.44 × 0.44 μm^2 per pixel).

Recording behavior in freely moving animals

The fish at 5–6 dpf with strong behavioral responses to moving gratings were selected on the days of OKR or OMR experiments. To determine the optomotor responses of freely swimming larvae to optic flow stimuli, we built a cylindrical swimming reservoir (diameter: 20mm, depth: 3mm) on the centers of a 50mm-diameter round transparent plastic plate. To make the stimulus displayed on the LED arena completely visible, the cylindrical reservoir wall was made from 2% low melting-temperature agarose (i.e. no plastic wall) to prevent plastic-agarose interfaces, which would introduce additional optical artifacts. This reservoir was placed in the centers of the LED arena, whose radius was >9 times larger than the radius of the reservoir. Because the optic flow stimuli were all symmetric to the horizontal plane, we do not expect any strong disturbance from internal reflection to animal behavior as discussed in a previous paper.[48] A selected fish was then placed in the reservoir filled with E3 medium.

An infrared LED lamp (850 nm, Conrad, Item No. 491248-62) was placed above the LED arena to provide relative homogeneous light for illumination. A CMOS camera (DMK23UV024, The Imaging Source GmbH, Bremen, Germany) with a 6 mm lens (C-Mount Lens FL-HC0614-2M, Ricoh) was placed at 30 mm below the reservoir. Fish behavior was recorded in 100 frame-per-second (fps) with a custom written Python application.

The body axis, body position, and raw eye angles were extracted offline by a custom written OpenCV based Python script. Briefly, the eye pixels and body pixels were extracted using a threshold-based segmentation method. The eye and body centers positions were computed as the first moment of the corresponding pixel ensembles. To obtain the body axis or the head direction, we first applied the principal component analysis (PCA) to the body pixel ensemble to obtain the rough body orientation as the first principal component (PC) vector. The coarse body direction (the sign of the first PC vector) was then determined as the side with higher variation on the second PC axis. We further refined the body axis by computing the vector that is perpendicular to the line linking the centers of two eyes and pointing into the coarse body direction determined previously. The orientation of the two eyes was also computed with PCA in the same manner. The eye position was computed as the angular difference between the eye orientation and the body axis (Figure 6B). The eye traces were first smoothed with a moving mean filter with the window size of 25 frames (250 ms) to reduce noise interference. In order to extract the OKR-related slow phase eye movement, the eye angle changes faster than 90 degrees/second were set to zero. The total eye angle changes over the entire stimulus period were computed for the left and right eyes (Δleft and Δright).

Eye movement recording in mounted fish

The selected 5-6 dpf fish were immobilized in 1.7% low melting-temperature agarose (wt/vol, E3 medium) at the tip of a transparent plastic triangle stage (tip angle < 45%) on the experiment day. The agarose surrounding the eyes of the mounted fish were carefully removed to allow free eye movement. The triangle stage with the mounted fish was later transferred into a spherical glass container filled with E3 medium. The container was placed in the centers of the cylindrical LED arena. The body position and orientation of the

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

mounted fish was adjusted later, so it was in the centers of the LED arena with the dorsal side facing up (roll correction) and the head pointing to 0° elevation (pitch correction) and 0° azimuth (yaw correction).

A CMOS camera (DMK23UV024, The Imaging Source GmbH, Bremen, Germany) was placed at ~7 centimetres below the mounted fish (~3 cm below the bottom of the glass bulb) to record the horizontal eye movement. To create homogeneous light for eye illumination, a 3-cm-width white paper stripe was placed above the opening of the glass bulb to diffuse the light emitted from a high-power infrared LED light (850 nm, Conrad, Item No. 491248-62) placed behind the glass bulb. The eyes were recorded at 60 frames per second (fps). The eye movement recording and stimulus synchronization were achieved via the same custom written Python application for freely swimming fish recording. The eye tracking and pre-processing were done offline by a custom written Matlab script in the same way described in the section above.

Stimulus design for behavior experiments

Eight combinations of translational and rotational optic flow stimuli were used in the global motion decomposition experiments in freely swimming fish (Figure 6C), which includes the pure forward/backward translation stimuli, the pure CW/CCW rotational stimuli, and the 4 different mixtures of these pure translation/rotation stimuli. The design detail of these optic flow stimuli was described in a previous section ("Optic flow stimuli design"). The relative strength for the translation and rotation components in all mixed optic flow stimuli, w_{trans} and w_{rot}, were set to 1 and 0.67 respectively. All optic flow stimuli for OMR experiments were normalized by the maximum local movement speed to the range of 0-60 visual degrees/second. We also measured the baseline OMR and OKR activities as the swimming trajectory and eye movement when a static binary noise pattern was displayed.

As shown in Figure 6E, we designed 3 variants of translation stimuli to determine if and which visual features were involved in the conjugate eye movement observed in the optic flow decomposition experiment. The pure translation and "T+R" mixture optic flow stimuli in the OKR experiments in mounted fish are constructed in the exact same way as above, except the local motion speed in the optic flow stimuli was reduced to the range of 0-30 visual degrees/second to better trigger OKR behavior. The slanted translation was constructed by rotating the pure translation $\sin^{-1}\left(\frac{w_{rot}}{w_{trans}}\right) = \sin^{-1}(0.67) \approx 42.1°$, which corresponds to the azimuth of the contraction point of the "T+R" mixture. When the motion cue in one hemi visual field is closer than the other side, the speed difference is similar to the difference caused by the rotation interference. To mimic this, the asymmetric translation stimuli was constructed by multiplying each flow vector in the pure translation stimuli by the inverse of a distance factor (D) which was calculated as:

$$\boldsymbol{D} = \sqrt{1 + \left(\frac{w_{rot}}{w_{trans}}\right)^2 - 2\frac{w_{rot}}{w_{trans}}\cos(\angle \overrightarrow{P}_{azi})}$$

$\angle \overrightarrow{P}_{azi}$ is the azimuth angle of the unit vector $\overrightarrow{P}$ pointing from the observation point (the origin) to the location P in the visual field. The transformation mimicked the changes caused by shifting the observation points which are equidistant from all motion cues ($D = 1$ a.u.) in the pure translation stimuli to the left by $\frac{w_{rot}}{w_{trans}}$ a.u.

RETROGRADE LABELLING OF nMLF

The retrograde labelling of the nucleus of the medial longitudinal fasciculus (nMLF) in nls-GCaMP6s zebrafish larvae was performed as described previously.[49] Briefly, a 50% Texas Red dextran (10000MW, Thermo Fisher) in Evans solution was injected into the spinal cord of 4 dpf. larvae embedded in 1.7% agarose and anesthetized with tricaine. The injected fish was imaged with the two-photon microscope at 5 dpf.

QUANTIFICATION AND STATISTICAL ANALYSIS

Calcium data pre-processing

The midbrain and diencephalon of mounted fish were sampled in the calcium imaging recording in dorsoventral direction from 20 μm to 120 μm below the top of tectum with the step size 10 μm. The motion artifacts along the recording plane (XY plane) were corrected by a phase-correlation algorithm, and all visible neuron somata in the time-averaged image of the corrected recording video were automatically selected as ROIs by a marker-controlled watershed algorithm. Per-ROI calcium time series were extracted as the sum of all pixel values within the ROI for each frame.

A 3D z-stack that imaged along the dorsal-ventral axis from the top of tectum to deep ventral pretectum and dorsal thalamus (step size: 0.44 μm) was recorded for each fish. The ROIs in each recording were first registered to the corresponding z-stack. The neurons recorded in different fish were registered to the z stack of *Tg(elavl3:Hsa.H2B-GCaMP6s)mpn400* transgenic line provided by the Max Planck zebrafish brain atlas (mapzebrain) by registering the ROIs in each recording to the corresponding z-stack aligned with the atlas z-stack with a custom-written multimodal registration Matlab program.

High throughput motion RF estimation

The high throughput motion RF estimation method used in our study was described in detail in a previous study.[26] Briefly, the neural responses to the CMN stimulus were recorded with *in vivo* calcium imaging. As the abrupt, high amplitude increases of somatic

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

calcium signals are usually correlated with significant increases of neural firing rate, these events were detected in the calcium traces and extracted for each ROI to obtain a binary event train (0 = no event, 1 = event detected). ROIs with event trains that only showed sporadic activity (#events/#stimulus frames < 8 %) were excluded for the following steps as we found only few of them could yield successful RF estimations with this method.

For each event train, an event triggered average (ETA) was computed as the raw RF estimation as follows:

$$ETA = \frac{1}{n_{event}} \sum_{i=1}^{I} e_i B\left(\vec{V}_i\right)$$

$$n_{event} = \sum_{i}^{I} e_i$$

The e_i was the event state (0 or 1) for the i^{th} stimulus frame, and I is the total number of frames in the stimulus. To avoid the ambiguity in the vector summation, the corresponding motion vector field of the i^{th} frame $\vec{V}_i$ was transformed into a sparse, one-hot representation $B(\vec{V}_i)$ where the phase angles of motion vector are evenly divided into 16 direction bins as described in the previous study (Zhang and Arrenberg[26] for details). For instance, for a vector $\vec{v}$ in $\vec{V}_i$ whose phase angle is in the range of the k^{th} direction bin, its one hot transformation $B(\vec{v})$ will be a 16-element array where the k_{th} element is the vector norm $|\vec{v}|$ and the other elements are 0. Therefore, the size of these ETAs was 8 row×28 column×16 direction bins.

To identify the spatially connected ETA components significantly related with RF structures, a two-step non-parametrical cluster-based bootstrapping test (2-step NCB test) is applied. Briefly, the event train of each neuron was circularly permuted with random temporal offsets for 1000 times. By comparing the original ETA with the null ETAs computed from these permuted event trains, we determined an empirical probability for each unit in the original ETA. The original ETA units with an empirical probability value greater than 97.5% (their values are higher than 97.5% of the corresponding units in the null ETAs) were selected as potentially significant units (step I). To solve the multiple comparison problem in the detection of significant RF structures, we performed an additional cluster-based bootstrapping test (step II) where we grouped the selected ETA units in the original ETA connected in spatial and directional domain into significant clusters. We defined the cluster-level statistic as the sum of the empirical probability of all units in a significant cluster. By applying the same protocol to the null ETAs, we obtained a null cluster-level statistic distribution for each neuron. The clusters in the original ETA whose statistic were higher than 95% of the values in the corresponding null distribution are considered as significant RF components at the significance level of 5% (see the corresponding paper[26] for detailed explanations). For the sake of clarity and convenience, the identified significant ETA components were converted from the one-hot representation form back to the vector field form for the downstream analyses and visualization. An RF mode is defined as a cluster of the spatially connected ETA components. The centers and preferred directions of RF modes were calculated as the averaged location and preferred direction of the components in each mode.

Fitting optic flow fields to RFs

To determine the similarity of the estimated RF to translational or rotational optic flow fields, we identified the most similar translational and rotational optic flow field for each RF estimation by fitting the optic flow fields to RF estimations with the least squares estimation method. The residual sum of squared angular difference (RSS_{angle}) between an RF estimation ($\vec{E}$) and the fitted optic flow field ($\vec{F}$) were calculated as:

$$RSS_{angle} = \sum_{i=1}^{I} \angle\left(\vec{E}_i \vec{F}_i\right)^2$$

$$\angle\left(\vec{E}_i \vec{F}_i\right) = \arccos\left(\frac{\vec{E}_i \cdot \vec{F}_i}{\left|\vec{E}_i\right|\left|\vec{F}_i\right|}\right)$$

We solved the fitting problem by searching the translation/rotation axis that minimized the RSS_{angle} with the Matlab function "*fminunc*", a gradient descent optimizer. The similarity between the fitted translation/rotation flow fields and the motion RF estimations were calculated as the averaged cosine similarity between each vector pair of the RF-related ETA components and the optic flow fields as follows:

$$\text{Similarity} = \frac{1}{I} \sum_{i=1}^{I} \frac{\vec{E}_i \cdot \vec{F}_i}{\left|\vec{E}_i\right|\left|\vec{F}_i\right|}$$

In addition, the symmetry of bimodal translation RFs in relation to the preferred translation directions in Figure 4E was calculated as:

$$\text{Symmetry} = 1 - abs\left(\frac{2\angle\left(\vec{\Delta}_{center}, \vec{T}\right)}{\pi} - 1\right),$$

$$where\ \vec{\Delta}_{center} = \mu_{left} - \mu_{right}$$

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

The μ_{left} and μ_{right} are the left and right mode center positions of the bimodal RFs. The bimodal RF is most symmetric when the vector linking its mode centers is perpendicular to its preferred translation direction $\left(\angle(\vec{\Delta}_{center}, \vec{T}) = \frac{\pi}{2}\right)$.

Prediction of optic flow preference

To determine how well the responses of corresponding neurons to optic flow can be explained by their RFs, we recorded the calcium responses of neurons in optic tectum and the pretectal areas in 3 fish to 36 global optic flow stimuli which include 6 pure translational and 6 pure rotational optic flow and the 24 combinations of these translational and rotational flow in random order (Figure 3E). Due to the memory limitation of our stimulus setup, we were unable to test all the combinations of translation and rotation optic flow fields, the 12 untested combinations were colored in gray in Figure 3E. We chose these 12 combinations for exclusion, because their FoC/FoE positions are not shifted by the rotation components; we therefore expect that decomposition of rotation and translation is more easily accomplished for these combinations. The maximum speed for all optic flow stimuli were 30 degree/seconds. The translation and rotation component in the mixture optic flows were mixed in the ratio of 2:1 (w_{trans} and w_{rot}).

Each optic flow stimulus was displayed at 15 fps for 4 seconds followed with a 6-second pause before switching to next stimulus. The binary noise pattern of the last frame of the previous stimulus phase lasted over the pause period to become the first frame of the next stimulus phase, so the stimuli were switched seamlessly. After all optic flow stimuli had been displayed, the CMN stimulus was displayed for RF estimation in the same neurons.

The means and standard deviations of calcium fluorescence activities of each ROI in the 16-second period before and after all stimuli had been displayed were measured as μ_{init} / μ_{end} and σ_{init} / σ_{end}. The ROIs whose $|\mu_{init} - \mu_{end}| > 4\sigma_{init}$ and/or $\sigma_{init}/\sigma_{end} > 4$ were excluded as either they might be bleached or drifted out of the imaged region. The calcium activities of the remaining ROIs were converted into z-score calcium traces as follows:

$$\text{Z-score} = \frac{F - \mu_{init}}{\sigma_{init}}$$

The z-score calcium traces (2 fps) were upsampled to match with the stimulus frame rate (15 fps). The response intensity (R) for each stimulus phase was calculated as:

$$R = \max(R_{max} - R_{init}, 0)$$

R_{max} is the maximum normalized calcium activity (z-score) in a stimulus phase and R_{init} corresponds to the normalized activity in the first frame of the stimulus phase (same as the last frame of the pause). Only the ROIs with estimated RFs[26] and strong responses to optic flow stimuli ($max(R) > 3$) were selected (1139 out of the 1484 ROIs with estimated RF in the experiments for Figures 3E–3G). The RF-predicted responses ($R_{prediction}$) to the presented optic flow stimuli ($\overrightarrow{GF}$) were calculated as follows:

$$R_{\text{prediction}} = max\left(\frac{1}{N}\sum_{n=1}^{N}\frac{\overrightarrow{RF_n}\cdot\overrightarrow{GF_n}}{\left|\overrightarrow{RF_n}\right|\left|\overrightarrow{GF_n}\right|}, 0\right)$$

N is the number of significant components in the RF estimations, $\overrightarrow{RF_n}$ and $\overrightarrow{GF_n}$ are the vectors of the n^{th} components in the RF and in the optic flow field respectively. Note in the following section, this function will be addressed as the ''RF-based response function''.

The translation-rotation index (TR index) was computed as:

$$\text{TR index} = \frac{Rmax_{trans} - Rmax_{rot}}{Rmax_{trans} + Rmax_{rot}}$$

$Rmax_{trans}$ and $Rmax_{rot}$ are the maximum response intensity of all pure translation or rotation optic flow stimuli respectively.

Translation direction and speed tuning

To determine the tuning properties of the translation-sensitive neurons found in the pretectal area, we measured their responses to the translational optic flows moving in 8 horizontal directions (azimuth: 0°,45°,90°,135°,180°,225°,270°,315°; elevation: 0°, Figure S4A) and 5 speed levels (7.5, 15, 30, 60, 120 degree/seconds). Each stimulus phase lasted for 15 seconds, including a 5-second moving phase and a 10-second pause phase. The stimuli were displayed in random order to avoid any historical effect.

The data pre-processing and the calculation of response intensity were the same as described in the previous section. We employed the discrimination index (DI) used in a previous study as a quantitative measure for the direction and speed tuning intensity separately:[50]

$$\text{Discrimination Index} = \frac{R_{max} - R_{min}}{R_{max} - R_{min} + 2\sqrt{SSE/(N-M)}},$$

where R_{max} and R_{min} are the responses to the most and least effective stimulus, and the SSE is the sum squared error of the averaged responses, N and M are the number of trials and the degrees of freedom, respectively. Because the DI for one parameter of interest may be affected by the other irrelevant parameter, the R_{max} and R_{min} were calculated as the maximum and the minimum responses averaged for the irrelevant stimulus parameter. This way we could quantify DIs separately for speed and direction.

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

To quantitatively summarize the direction and speed tuning properties for all translation optic flow sensitive neurons, we fitted the joint distributions of von Mises and log Gaussian distributions to the neuron response data with the nonlinear optimization function "*fminsearch*" *in* MATLAB. The fitted distribution was:

$$\text{von Mises-log Gaussian distribution} = \overbrace{\frac{1}{v\sigma\sqrt{2\pi}}\exp - \left(\frac{\ln(v - \mu_s)^2}{2\sigma^2}\right)}^{\text{Log Gaussian term}} \overbrace{\frac{e^{\kappa\cos(\theta - \mu_d)}}{2\pi I_0(\kappa)}}^{\text{von Mises term}}$$

The log Gaussian and the von Mises term aimed to capture the speed and direction tuning properties, respectively. The preferred speed and direction of each neuron can therefore be quantified as the mode of the log Gaussian model $e^{\mu_s - \sigma^2}$ and the von Mises model (μ_d) respectively. The measure of concentration of the von Mises distribution κ provides a quantitative measure of the direction tuning width. Note that the choice of log Gaussian function was mainly due to the non-negative property of speed and the similar shape to the speed tuning curves observed empirically. There was no biological basis for this model thus it was used only for description purpose.

Population decoder for self-motion estimation

To determine if the information encoded by the translation-sensitive neurons was sufficient for determining the translation direction and speed, we constructed a population decoder derived from the maximum likelihood decoder with the calcium response data recorded. Since the accuracy of any decoding algorithm is limited under the Cramér Rao bound which is determined by the encoding information,[51] a low or no decoding error may prove the information encoded by these translation sensitive neurons is sufficient, even if the decoding algorithm implemented in zebrafish brain might be very different from our decoder.

As illustrated in Figure 5A (top), the ground truth stimulus parameter T is estimated as the least squares fit that minimizes the residual sum of squares between the calcium responses R of the recorded n translation-sensitive neurons, and the responses predicted by their response functions $(F(T))$:

$$\hat{T} = argmin\left(\sum_{i}^{n}(F_i(T) - R_i)^2\right)$$

We computed the RF estimation of the neural responses to 1000 evenly sampled translation directions. The least squares fit, or the arguments of the minimum (argmin) is approximated by finding the nearest neighbour of the recorded responses in the sets of estimated responses with the Matlab function "*knnsearch*". The histogram in Figure 5A contains 300 data points for "translation only," which correspond to the six translation directions for which we had ground truth neuronal responses and 50 different bootstrapped decoders. Each decoder was constructed by drawing 50 random RFs out of the 138 translation RFs characterized in Figure 3E. The "mixture" histogram is based on 24x50 datapoints (24 global flow combinations from data in Figure 3E).

The responses functions were approximated with either the RF-based response ($R_{prediction}$) function (Figure 5A, bottom, see the STAR Methods section "prediction of optic flow preference" for details) or the joint distribution model fitted to the direction-speed tuning data (Figures S4B–S4E, S5A, and S5B). Since the decoding accuracy was limited by the encoding quality, the decoding error provides a quantitative measure for the encoding quality of translational optic flow. The error was calculated as the absolute difference between the real stimulus parameter and the estimation. In the decoder for translation and rotation mixtures (Figure 5A, bottom), the RF-based predictions to optic flow (from RF estimation using the CMN stimulus) were compared to the measured response to a particular optic flow (data from Figure 3E).

One problem in estimating the encoding quality of translation direction and speed using the joint distribution model in Figure S4B, is that the tuning model already contains the response information of the target stimuli to be tested with (which is different from the situation for the model at the bottom of Figure 5A). We employed a jack-knife test ("leave-ten-out") to avoid this circular analysis: in the decoder for Figures S5A and S5B, the neural response data to translational optic flow with 40 different combinations of direction and speed parameters were evenly divided into 4 groups, so each group contained the responses to 10 different optic flow stimuli. For any target stimulus to be decoded, the decoder was only allowed to access the groups that did not contain the responses to the target stimulus for tuning model estimation.

To test the robustness of translation direction encoding algorithms implemented with the translation RFs, we constructed a population autoencoder which encodes the partially occluded or tampered translational optic flow field (used in Figures 5B and S5E) with n translation RFs into a one-dimensional array of neural responses as:

$$R_j = F_j(S) + \varepsilon,$$

$$where\ F(S) = max\left(\frac{1}{I}\sum_{i=1}^{I}\frac{\overrightarrow{RF_i}\cdot\overrightarrow{GF_i}}{\left|\overrightarrow{RF_i}\right|\left|\overrightarrow{GF_i}\right|}, 0\right)$$

Please cite this article in press as: Zhang et al., A robust receptive field code for optic flow detection and decomposition during self-motion, Current Biology (2022), https://doi.org/10.1016/j.cub.2022.04.048

R_j and $F_j(S)$ are the simulated recorded response and the RF-based response function of the j^{th} simulated neurons to stimulus S, and ε is the noise term. The neural responses were decoded to estimate the translation direction with the same population decoder described above, which makes the decoding process basically a reverse operation of the encoding. Thus, the reconstructed error is representing mostly the encoding quality which is affected by the number of neurons involved (n), the noise in the encoding process (ε), and the input translational optic flow field $(\overrightarrow{GF})$.

To determine the noise effect on the encoding process, the reconstruction error of the autoencoder using all 329 translation RFs was tested with different encoding signal-to-noise ratio (SNR = $+\infty$ (ε = 0), 2.6, 1.3 and 0.65) which was calculated as:

$$SNR = \frac{\mu}{\sigma},$$

The μ is the averaged output from the response functions of all 329 simulated neurons, and σ is the standard deviation of the noise ε. For each SNR, the reconstruction error was tested for 100 random translation directions (Figure S5C).

We also evaluated the encoding quality for 100 random translation directions when the number of translation RFs involved n = 10, 20, 50, 100 and 329 (Figure S5D, SNR = 1.3). Based on the results in Figures S5C and S5D, the autoencoder used in Figures 5B and S5E was constructed with 50 translation RFs and the SNR = 1.3. As a reference, on average, more than 100 translation sensitive neurons were found in the pretectal area per fish in the *in vivo* calcium imaging experiments, and as mentioned previously, the calcium signal SNR of the optic flow sensitive neurons analysed were all higher than 3.

The optic flow in naturalistic scenes is contiguous, so are the occlusions and object motions that partially block or tamper the optic flow fields. To mimic this contiguity, we created a 2D Gaussian blurred white noise matrix (σ = 12 visual degree) which had the same size as the optic flow fields for each optic flow field. Only the corresponding vectors in the optic flow field of the top N% elements (sorted by value) in the noise matrix were preserved in the final flow field (examples in Figure S5F). The rest of the optic flow field were occluded or replaced with random object motion (Figures 5B and S5E). For the latter, each spatially connected region in the rest of the optic flow field was assigned with a single random object motion. Similar to the calculation of translational optic flow fields, let the unit vector $\overrightarrow{O}$ indicates the object motion in the 3D environment, the motion vector $\overrightarrow{V}$ in a region of the visual field indicated by a unit vector $\overrightarrow{P}$ was calculated as:

$$\overrightarrow{V} = \overrightarrow{P} - \left(\overrightarrow{P} \cdot \overrightarrow{O}\right)\overrightarrow{O}$$

Assessing vergence of optokinetic responses

A conjugate-vergence index (CV index) was computed for evaluating whether the OKR to each stimulus was more similar to a conjugate movement or a vergence movement:

$$\text{CV index} = \frac{1}{2\pi}\arctan\left(\left|\frac{\Delta_{right} - \Delta_{left}}{\Delta_{right} + \Delta_{left}}\right|\right)$$

The CV index equals to 0 and 1 when the fish perform pure conjugate and vergence eye movement respectively. The CV indices of OKR to different optic flow stimuli were compared with one-sided, Bonferroni-corrected Wilcoxon rank-sum test to determine if any of them are significantly more similar to conjugate eye movement.

To determine if the conjugate eye movement to the rotation component in the T+R mixture optic flow is caused by local statistics, we randomly sampled the hemi optic flow fields for 1000 times and computed the averaged motion vectors for the left and right hemi field separately. The bootstrapped distribution of CV index in Figure 6F was computed in the same way from these pairs of bootstrapped motion vectors as described above.

Current Biology, Volume *32*

Supplemental Information

A robust receptive field code for optic flow detection and decomposition during self-motion

Yue Zhang, Ruoyu Huang, Wiebke Nörenberg, and Aristides B. Arrenberg

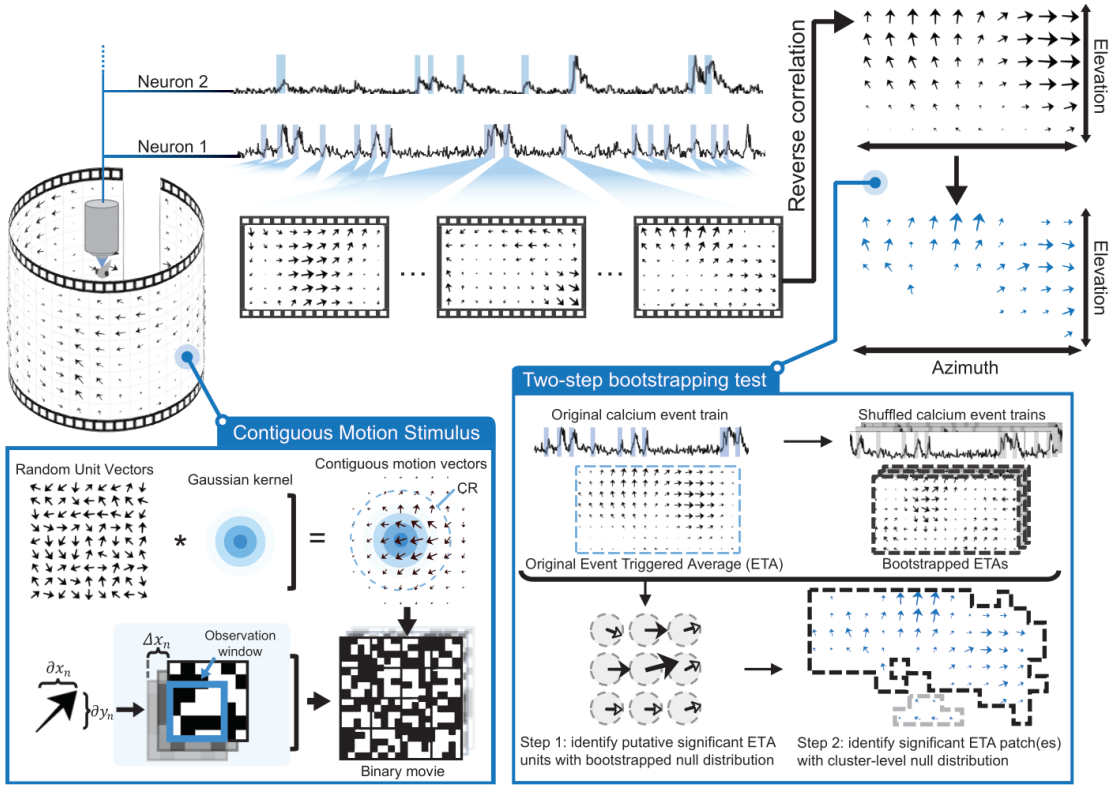


Figure S1. A graphical abstract for the receptive field estimation method using contiguous motion stimulus and two-step bootstrapping test (adapted from Zhang and Arrenberg, 2019, see STAR Methods for details). Related to Figure 1B and STAR methods.

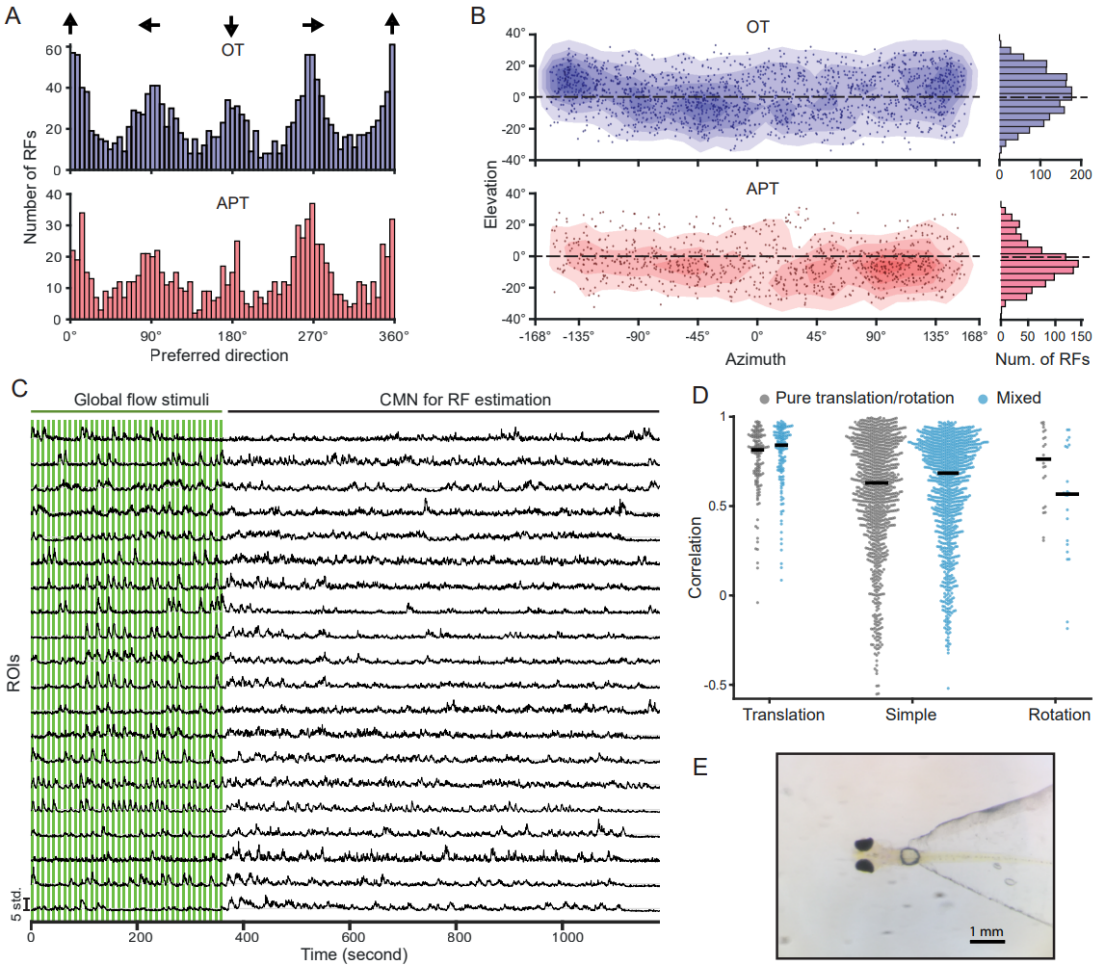


Figure S2. Unimodal RF characterization and quantitative evaluation of complex RFs, related to Figure 1, 3 and STAR methods. The distribution of averaged preferred directions (A) and RF centre locations (B) in the visual fields for the neurons with unimodal RFs in optic tectum (OT) and area pretectalis (APT). The black arrows in the top row of A indicate the preferred motion direction in 2D. Each dot in (B) indicates the centre of an RF on the 2D spherical coordinate map. The filled contours in the back of the scatter plot indicate the RF centre density estimated with kernel density estimation. The histograms on the right show the distribution of RF centre elevation. The RF centres of most unimodal neurons (66.8 %) in the APT but not the OT (45.8 %) were in the lower half of the visual field. (C) Examples of z-score calcium traces recorded in the experiment for Figure 3E-G. Each green shaded area indicates an optic flow stimulus phase (4 seconds duration + 6 seconds pause interval). (D) The bee-swarm plots of the correlation between the predicted and real neural response intensity to pure translational/rotational or mixed optic flows used in Figure 3E. The neurons were grouped by their RF types indicated in the x axis. (E) An image of an agarose-mounted 5 dpf zebrafish larva on the tip of a plastic triangle stage as used in this study.

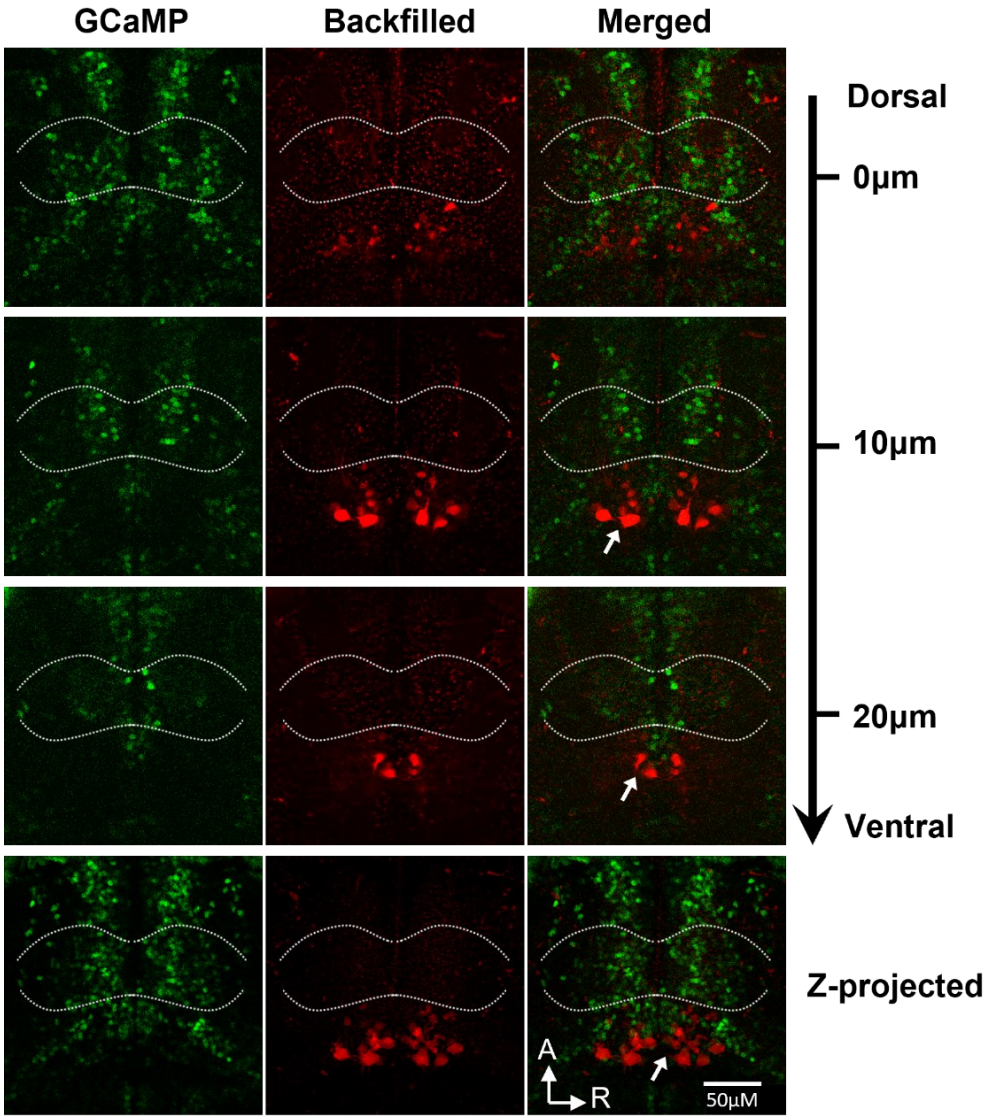


Figure S3. Dorsal view of the pretectal area in 3 different depths (related to Figure 4B). The green channel (GCaMP) displays all visible neural somata (two-photon calcium image) in the corresponding area, the white dashed curves outline the approximate location of the AMC pretectal clusters. The red channel shows the neural structures retrogradely labelled by spinal cord injections (see also STAR methods “Retrograde labelling of nMLF”). The white arrow in the merged channel indicates the nucleus of the medial longitudinal fasciculus (nMLF). The bottom rows contain the superimposed images of the pretectal area in different depth (Z-projected); A: anterior direction; R: right direction.

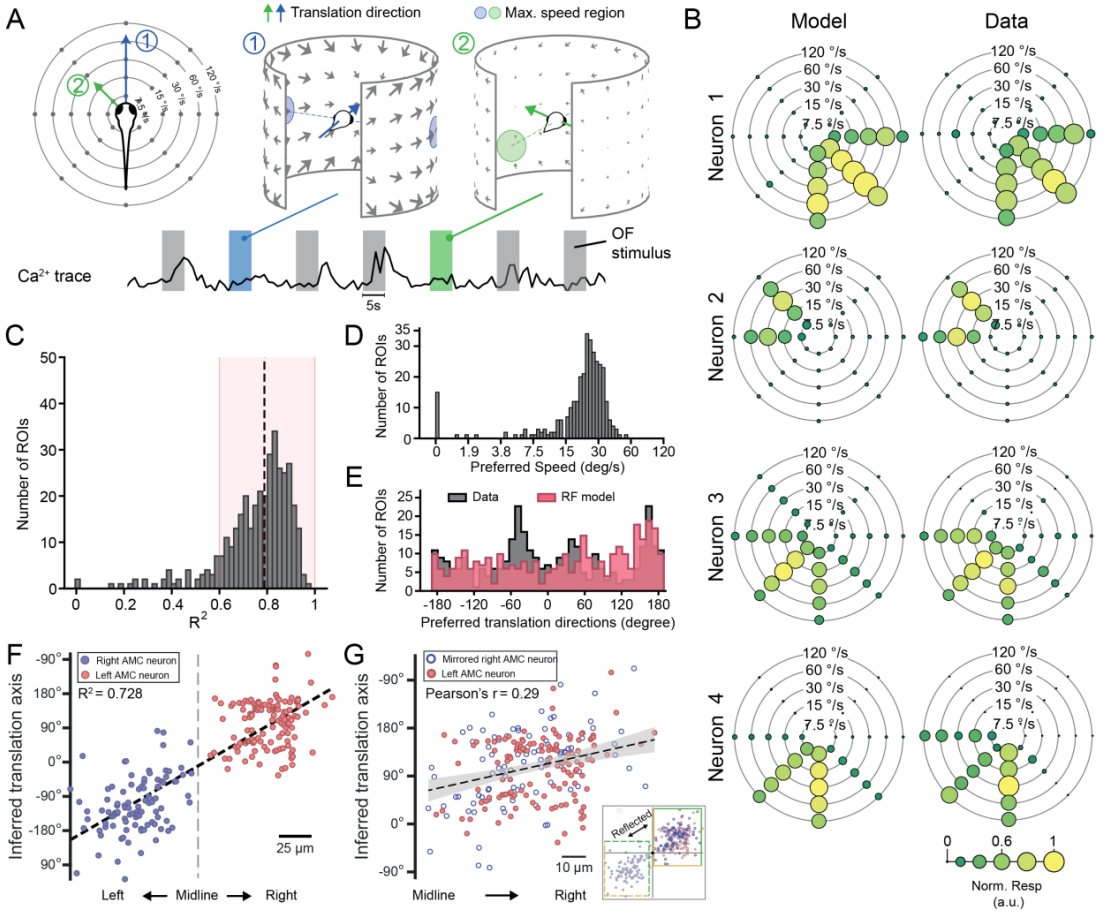


Figure S4. Joint encoding of translational optic flow direction and speed in the pretectum, related to Figure 4. (A) Left: Animals were presented with stimuli consisting of one of eight different TOF directions and five logarithmically spaced maximum velocities. Right: Two examples TOF stimuli are shown (① and ②). The region of maximum velocity of each stimulus is indicated as grey disk and located to the side of the translation direction (black arrow). To avoid the speed/distance ambiguity in TOF, which is not the focus of our study, all visual cues in our OF stimuli were equidistant to the animal. Therefore, the maximal angular speed in the translational optic flow is proportional to the fictive translation speed of the animal.(B) The recorded response (right column) and the fitted tuning model (left column, joint model of von-Mises and lognormal distributions)) for 4 example neurons. The response intensity is normalized by the averaged response amplitude for visualization purposes. (C) The R^2 distribution for the tuning model fitted to the neural responses to translational optic flow with different speed and/or direction. The black dashed line indicates the median R^2 and the red shaded area indicates the ROIs (R^2>0.6) used in (D-F) and Figure 4F. (D) The preferred translation speeds of the translation-sensitive neurons inferred from the fitted tuning map. (E) The distribution of the preferred horizontal translation direction inferred from the fitted tuning model ("Data") and from the RFs of the neurons in the same anatomical locations ("RF model"). (F) Linear regression (black dashed line, R^2 = 0.728) between the preferred translation directions and the locations of AMC neurons on the left-right body axis. The dashed line indicates the midline of fish. (G) Intra-cluster linear correlation (Pearson's r = 0.29, $p < 0.001$) between preferred translation directions and the anatomical locations. The right AMC dots in (F) were reflected in the origin.

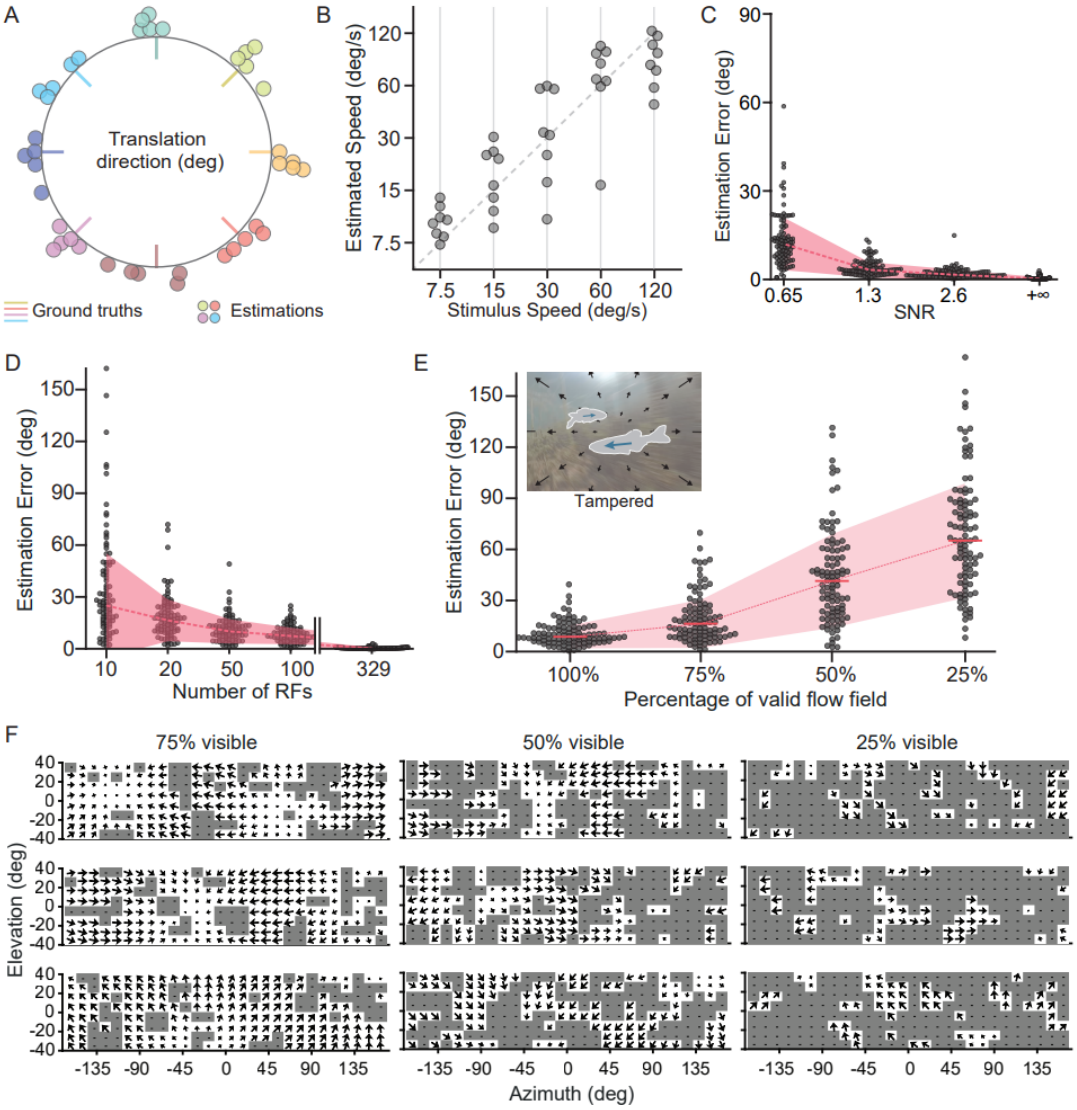


Figure S5. Encoding accuracy of self-translation information in the population code of translation RF neurons, related to Figure 4, 5 and STAR methods. (A) The translation direction estimated with the data presented in Figures 4F and S4A-E plotted in bee swarm style. Coloured bars indicate ground truth horizontal translation directions and circles the estimations. (B) The translation speed estimated with the data presented in Figures 4F and S4A-E in bee swarm style. The dashed line indicates the line of unity. (C-D) The bee-swarm plots of the estimation errors in relation to the signal-to-noise ratio (SNR) in the optic flow vector field (C) and the number of RFs contributing to the autoencoder (D). (E) Bee swarm plots of the estimation errors for the TOFs corrupted by moving objects. The dashed lines and shaded area in (C-E) indicate the median ± the standard deviation. (F) Examples of incomplete or tampered translational optic flow fields. Each grey block indicates a region in the visual field that contains no or wrong motion cue. In the tampered visual field, each set of spatially connected grey blocks displayed the 2D projected motion resulting from the 3D movement of a single object covering this region.

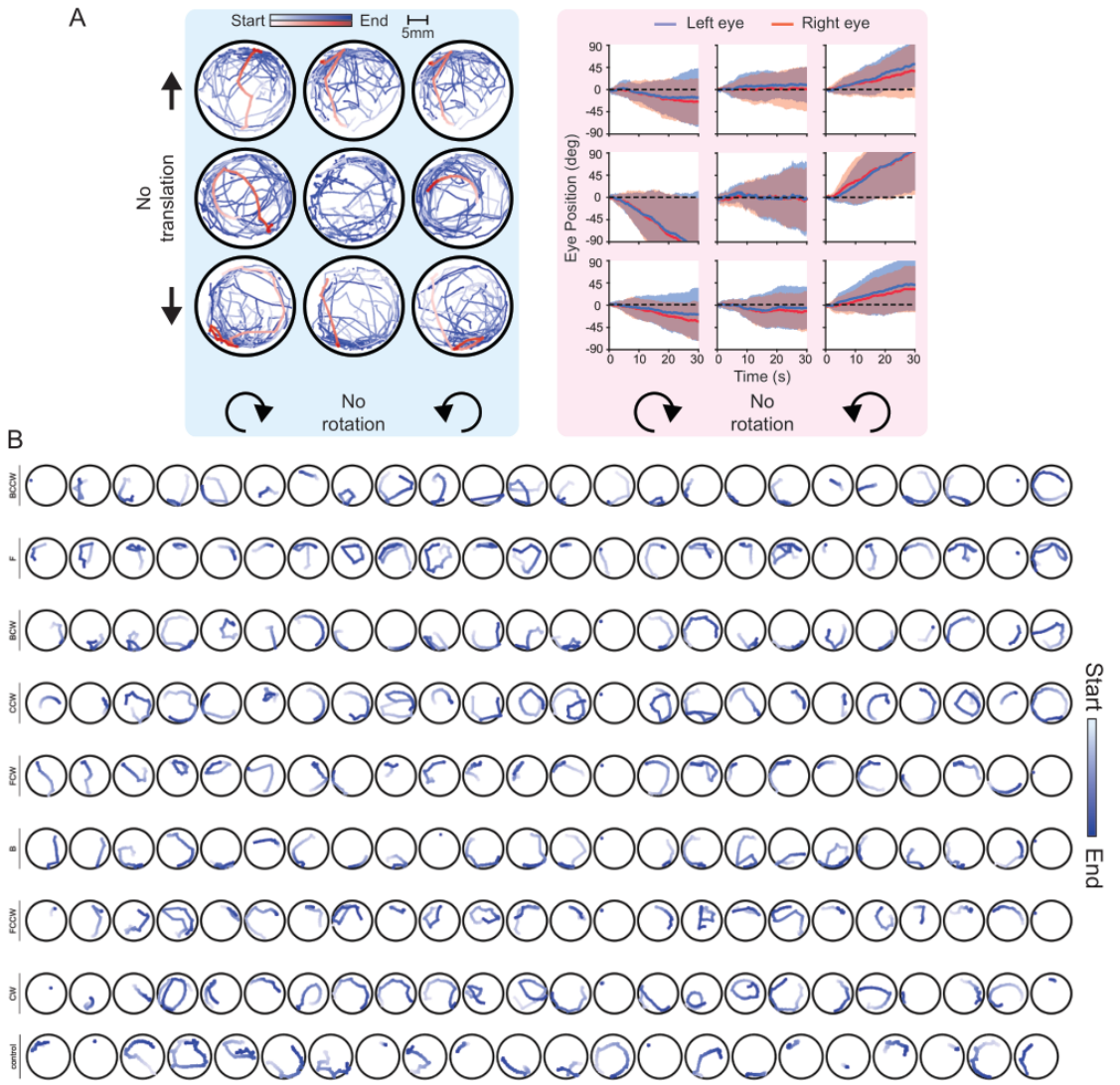


Figure S6. OKR and OMR responses to all optic flow stimuli used (related to Figure 6). (A) The superimposed swimming trajectories (left) during all 9 optic flow stimuli phases (indicated by the arrows at the sides) and the corresponding eye position traces (mean ± standard deviation). (B) The individual swimming trajectories for each optic flow stimulus (rows) and fish (column). From top to bottom: BCCW (backward translation + counter-clockwise rotation), F (forward translation), BCW (backward translation + clockwise rotation), CCW (counter-clockwise rotation), FCW (forward translation and clockwise rotation), B (backward translation), FCCW (forward translation + counter-clockwise rotation), CW (clockwise rotation) and control (stationary stimulus).