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

Authors/affiliations

Zhang, Yue[1,2]; Huang, Ruoyu[1]; Nörenberg, Wiebke[1,2,3]; Arrenberg, Aristides[1]

Author list footnotes

1. Werner Reichardt Centre for Integrative Neuroscience, Institute of Neurobiology, University of Tübingen, D-72076, Tübingen, Germany
2. Graduate Training Centre for Neuroscience, University of Tübingen, D-72076, Tübingen, Germany
3. Present address: Humboldt-Universität zu Berlin, Faculty of Philosophy, Berlin School of Mind and Brain, Germany; Humboldt-Universität zu Berlin, Faculty of Life Sciences, Department of Psychology, Germany.

Contact info

aristides.arrenberg@uni-tuebingen.de

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 behavioural 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.

Keywords: diencephalon, self-motion, optic flow decomposition, zebrafish, receptive field, visually guided behaviour, optokinetic response, optomotor response, matched filter algorithm

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 (Gibson, 1954; Lappe et al., 1998; Nakayama and Loomis, 1974). Optic flow is used for self-motion estimation (Britten, 2008; Festl et al., 2012; Jr et al., 2001; Srinivasan et al., 1996; Warren and Hannon, 1988), which is a critical task for all visually guided navigation, as well as for gaze and posture stabilization behaviours.

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 (Duffy and Wurtz, 1991, 1995; Luo et al., 2019; Sunkara et al., 2016; Tanaka and Saito, 1989; Warren and Hannon, 1988), the nucleus lentiformis mesencephali in pigeons (Wylie, 2000), and the lobula plate in flies (Krapp and Hengstenberg, 1996). In zebrafish, the pretectal area processes optic flow and is necessary for body position and gaze stabilization in optomotor responses (OMR) and optokinetic responses (OKR), respectively (Kubo et al., 2014; Naumann et al., 2016; Wang et al., 2019, 2020).

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 (Dahmen et al., 2001; Koenderink and van Doorn, 1987; Longuet-Higgins and Prazdny, 1980). 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 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 (Dahmen et al., 2001; Franz and Krapp, 2000).

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 (Huston and Krapp, 2008; Krapp and Hengstenberg, 1996; Wertz et al., 2009). 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 (Krapp and Hengstenberg, 1996; Kubo et al., 2014). 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

recordings 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 stimuli and reverse correlation techniques developed recently for the systematic RF characterization of motion-sensitive neurons via calcium imaging (Zhang and Arrenberg, 2019). We measured the fine-grained motion RFs of thousands of neurons in the pretectum and optic tectum of larval zebrafish. The identified RF structures reveal a neural mechanism for robust optic flow encoding in the pretectum. Furthermore, we show that larval zebrafish decompose translation and rotation components in optokinetic and optomotor behaviour as suggested by the underlying encoding mechanisms.

Results

Identification of unimodal and bimodal receptive fields sensitive to motion stimuli

To estimate motion receptive fields (RFs), we used a full-surround cylindrical stimulus arena (Figure 1A) and recorded neuronal responses to motion stimuli in zebrafish larvae expressing the calcium indicator GCaMP6s. The stimulus consisted of a contiguous motion noise pattern in which moving directions were spatiotemporally coherent (Zhang and Arrenberg, 2019). Using reverse correlation, we estimated the RF by averaging across response-evoking stimulus frames and rejecting noise components based on a spatial and directional bootstrapping statistic (Figure 1A, also see Figure S1 and Methods for details). We successfully identified the RF structures of 3,926 motion-sensitive neurons in the optic tectum and the pretectal area of seven fish. The shape, scale, location, and direction preference of these RFs were highly diverse, as illustrated by the estimated RF structures of the four example neurons in Figure 1C.

Despite the high structural diversity, we found that most estimated RFs (99 %), contained either one or two significant modes (Figure 1B), 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 optic tectum (OT) and the area pretectalis (APT, Figure 1B,D), most unimodal RFs are located in the OT and the bimodal RFs are predominantly located in the APT (Figure 1B,D).

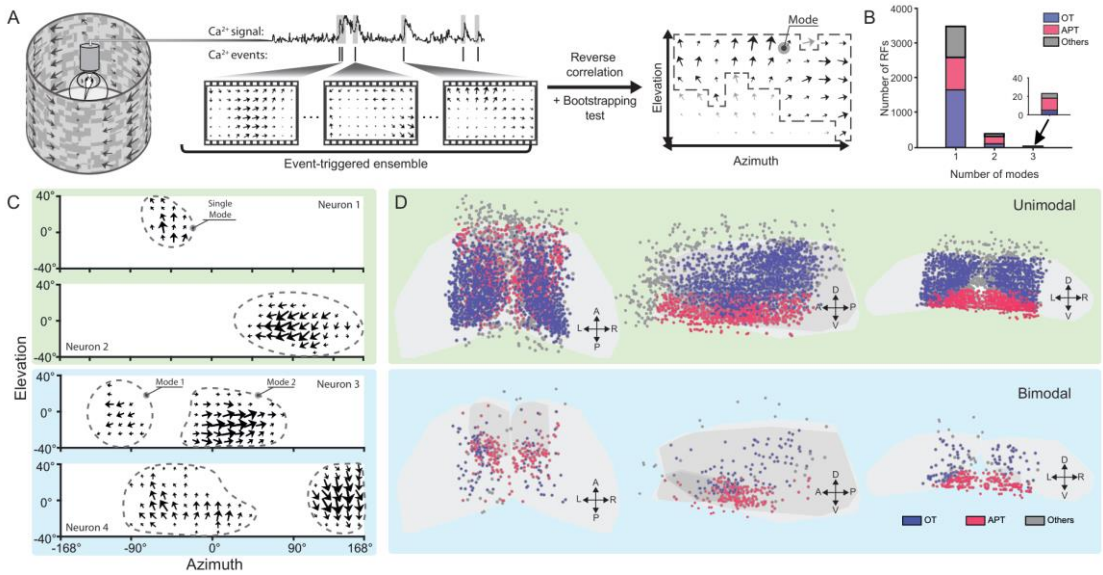

Figure 1 Characterization of unimodal receptive fields (RFs). (A) Illustration of the RF estimation method using a contiguous motion noise stimulus (see Methods). (B) Histograms of RF mode number for neurons in the optic tectum (OT), area pretectalis (APT) and adjacent brain regions. (C) Examples of 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. (D) Anatomical

locations of neurons with unimodal RFs (n=3504, in 7 fish) and bimodal RFs (n = 398). The anterior-posterior (AP), dorsal-ventral (DV) and left-right (LR) axes are labelled.

Asymmetric location and patterned direction selectivity of the modes in bimodal RFs

Due to the fundamental structural difference between unimodal and bimodal neurons, we characterized them in different ways. The averaged preferred direction of each unimodal RF corresponded to either up, down, left, or right (Figure S2), and the spatial RF distribution matched results of a previous study (Wang et al., 2020)

For the bimodal RFs, the preferred directions and spatial centres were calculated separately for each mode. To measure the span of bimodal RFs, we calculated the distance between the two mode centres in visual space (intermodal distance, illustrated in Figure 2A). The intermodal distances ranged from ca. 60° to 315° in our dataset (Figure 2B), and intermodal distances of 180° were scarce. During the RF estimation experiment, the eyes of the whole-mounted fish were in their normal resting position. According to a previous study (Bianco et al., 2011), the binocular overlap zone should be smaller than 30° in this condition as illustrated in Figure 2C. Therefore, the bimodal RFs spanning across two sides of the overlap zone are classified as binocular RF. Based on this principle and the intermodal distance, the bimodal RFs were divided into 3 groups: 1. Putative monocular RFs receiving visual input from the nasal (-45° to + 45 °) and temporal end (-168° to -90° or 90°to 168°) of a visual hemi-field (shown in red in Figure 2C & 2D); 2. Binocular RFs with short intermodal distance (average: ~150°, coloured in green); 3. Binocular RFs with long intermodal distance (average ~223°, coloured in blue). For most bimodal RFs, the mode centres were distributed asymmetrically in relation to the anterior-posterior (AP) body axis as shown in Figure 2C and 4E.

Next, we compared the preferred directions within the two modes of each bimodal RF by plotting the distribution of preferred directions in the left mode grouped by their right mode preferred directions (Figure 2D). We found that the RFs preferring leftward (clockwise from above) motion with their right modes are mostly preferring the opposite direction (rightward motion) in their left mode (Figure 2D), 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).

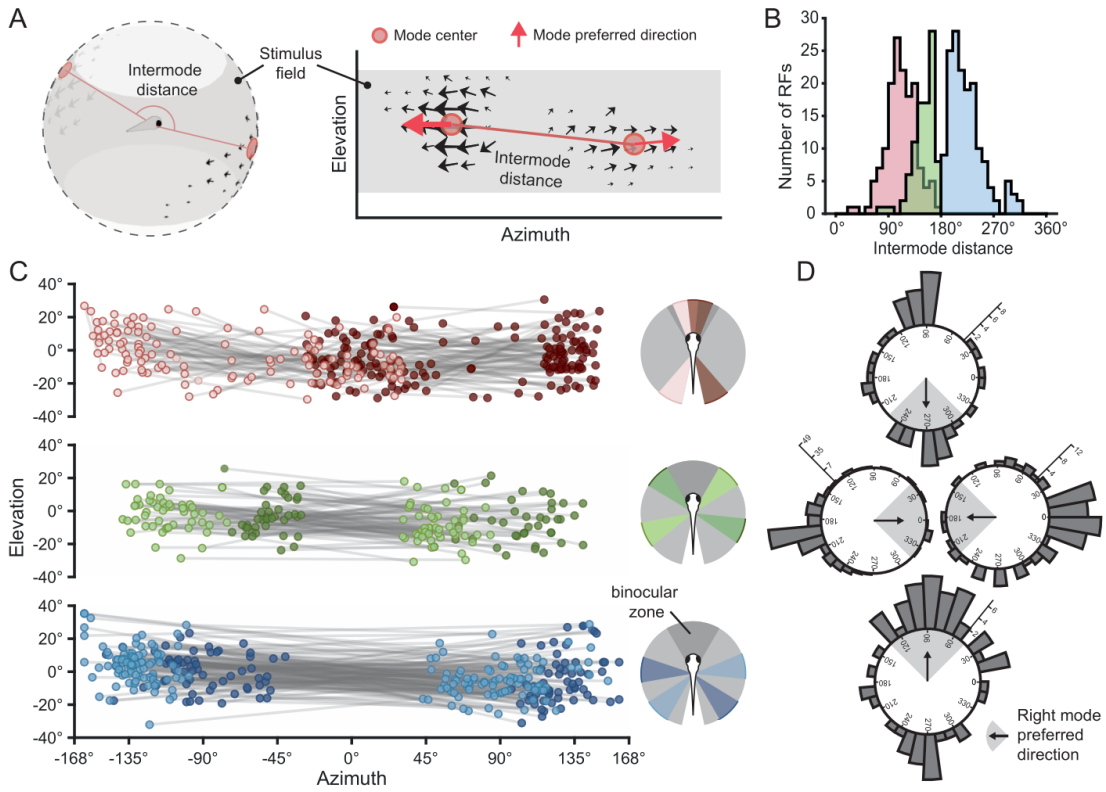

Figure 2 Spatio-directional organisation of bimodal RFs across the visual field. A) Illustration of a bimodal RF projected onto a sphere (left) and the corresponding 2D visual space representation (right). The "intermode distance" was defined as the distance between the two spatially separated components (modes) in the RF. (B- C) Bimodal RFs grouped by their inter-mode distance and locations in visual field. Each pair of linked dots in (C) indicates the mode centres of a bimodal RF. The RFs tilted to the right hemi-field were coloured 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).

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 motion in their left and right modes, respectively, are similar to the rotational optic flow field (ROF) 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). The similarity can vary from -1 to 1, where 1 means in the regions the RF covered, the RF structure is the same as the corresponding optic flow field, and 0 means the RF is not similar to any optic flow field.

For most RFs (89%), both highly similar TOFs and ROFs existed (cf. dots whose translation and rotation incongruity are both less than 0.2 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 (Figure 3B). These RFs were called “complex RFs” as they are mostly bimodal RFs (82%) containing heterogeneous vector field structures. In these complex RFs, the 329 RFs which were more similar to TOFs were called 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 (median cosine similarity = 0.98 and 0.92 respectively, Figure 3C). Only 23 of the 398 bimodal RFs were neither similar to any TOF nor ROF (the ones in the upper right corner of Figure 3B). These results suggest most bimodal RFs are similar to specific optic flow fields induced either by translation or rotational self-motion components.

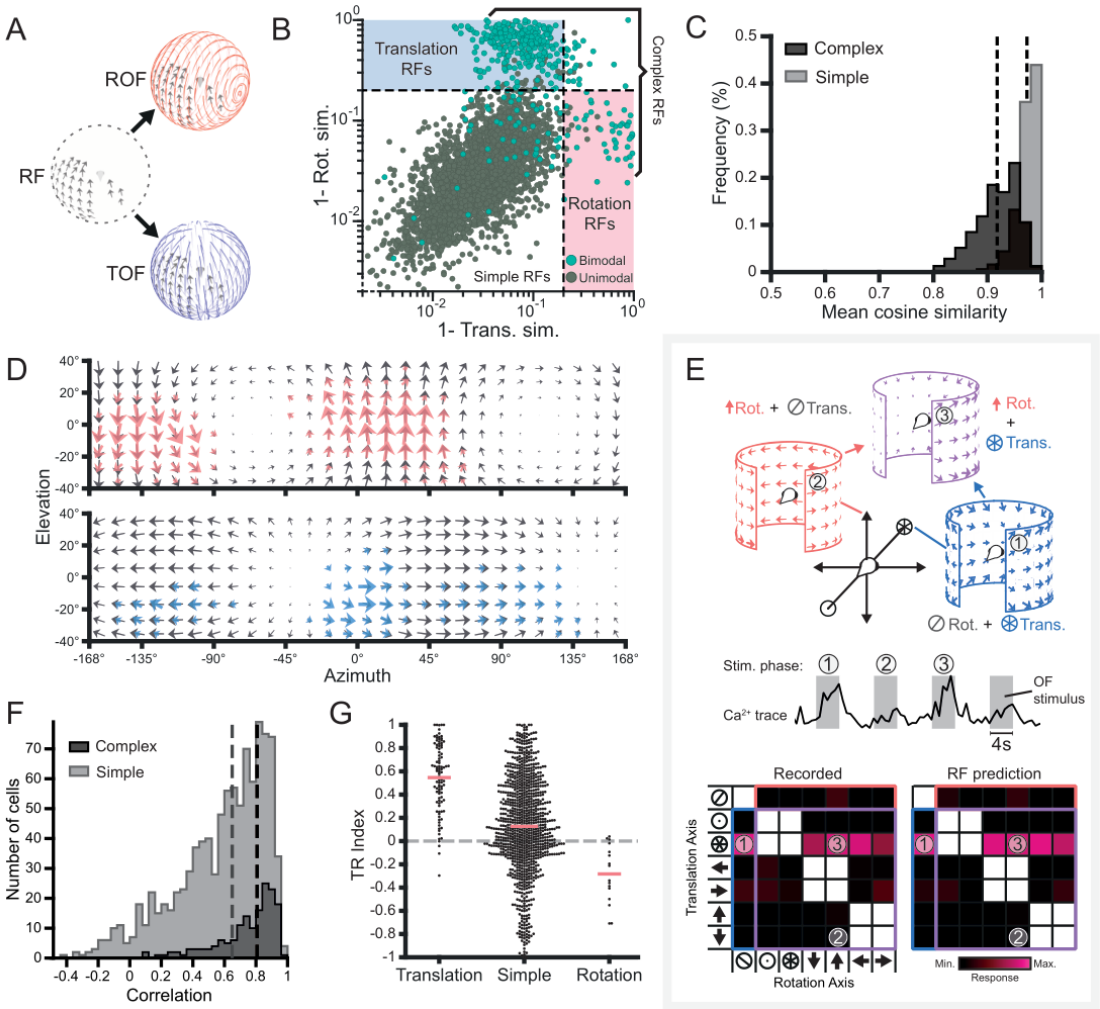

Figure 3 Prediction of optic flow responses in complex RF neurons with the matched filter model. (A) The best matching rotational (red) and translational (blue) optic flow field for an example motion RF (grey). (B) Score of the match of each RF to the fitted rotation (rot.) and translation (trans.) optic 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 (coloured) and the matching optic flow fields (grey); (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. (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 Methods) 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.

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 (Franz and Krapp, 2000). To test whether or not these complex RFs in larval zebrafish play a similar role, we displayed 36 distinct optic flow stimuli to the neurons whose RFs had been characterized earlier (Figure 3E and S3). Each stimulus consisted of either pure translation or rotation, or of a unique combination of translational and rotational optic flow. Using the measured RFs (Figure 1C), 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 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 suggests the complex RFs on average gave better prediction of the neural responses to optic flow stimuli than the simple RFs (Figure 3F and S4). We quantified the response specificity to translation vs. rotation (TR index, see 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 presentation of combined rotation/translation optic flow fields.

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.

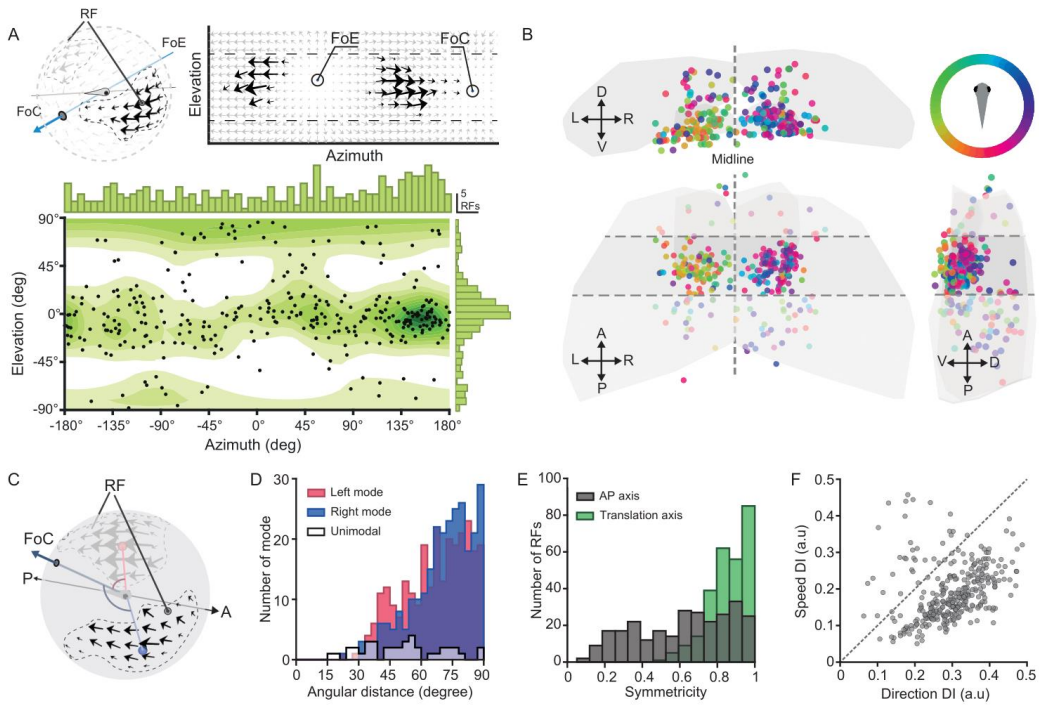

Figure 4 A topographic map of global translation direction in the pretectum. (A) Top: Illustration of focus of expansion (FoE) and contraction (FoC) of the fitted TOF field (grey arrows) of one translational RF (black arrows). Bottom: Preferred translation direction distributions for translation RFs (n=329). Each dot corresponds to the preferred translation direction (FoC) of a neuron. (B) Anatomical locations of neurons (n=329) with translational RFs, coloured according to preferred translation direction in the horizontal plane. The AMC pretectal region is contained within the horizontal dashed lines. (C) Illustration of the angular distance between mode centres and preferred translation direction (FoC) analysed in (D). The red and blue arcs indicated the angular distance of left and right mode centres respectively. (D-E). The angular distance of mode centres 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 to 120 °/s) and direction (0-360°) of the 369 translation-sensitive neurons in the AMC of two fish.

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 translational optic flows (Figure 4A).

The majority of these neurons were anatomically arranged in the anterior medial cluster (AMC) region (Kubo et al., 2014) of the APT (Figure 4B and S5). 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. Linear regression showed that the preferred translation direction is strongly correlated with anatomical position (R^2 = 0.72, Figure S6E). To visualize this anatomical arrangement further, we recorded the neural responses in the AMC to a TOF stimulus, whose direction changed clockwise (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 (Movie S1).

There are different ways in which a translation-sensitive neuron could detect the direction of translational optic flow. For example, its complex RF could sample the vicinity of the FoC and/or the 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 centres(s) and the preferred translation directions, including both unimodal and bimodal RFs in the analysis. We found that the mode centres of most bimodal translation RFs were located 60-90 degrees away from the preferred translation direction (Figure 4D), corresponding to regions with highest speed and strongest local motion coherence in the optic flow field. Furthermore, we found the mode centres of most bimodal translation RFs, which distributed asymmetrically in relation to the body axis (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 speed. Neurons were jointly tuned to translation direction and speed (Figure S6) and preferred speeds from approximately 15° - 40°/s. To estimate the fidelity of the joint speed and direction representation in the AMC neurons, we computed the discrimination index for translation direction and speed for each neuron. A higher discrimination index 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).

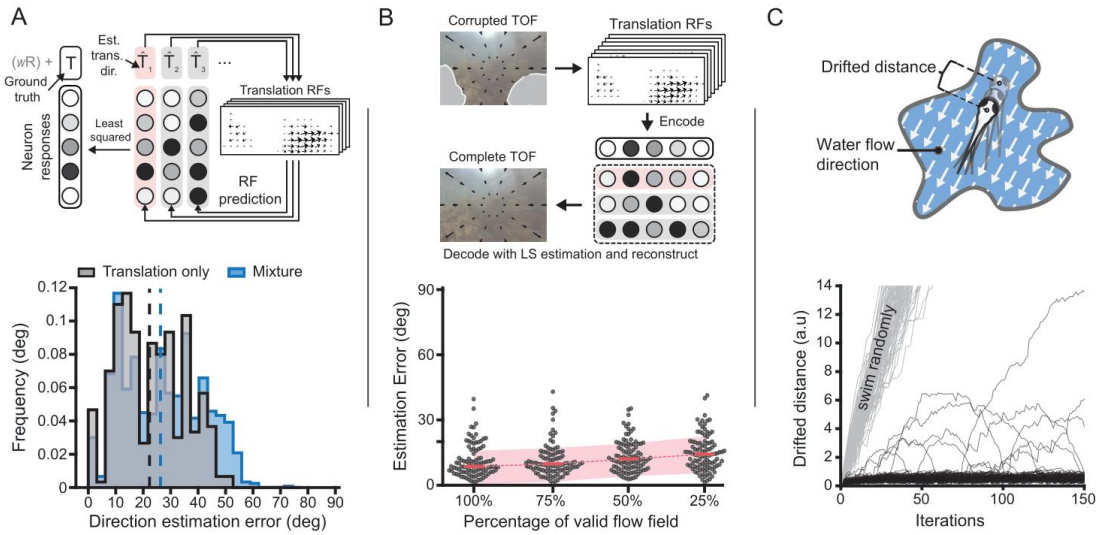

Figure 5 Robust encoding of translational optic flow direction by translation RFs. (A) Top: Illustration of the constructed population decoder estimating the stimulus parameters from the responses of optic flow sensitive neurons (see Methods section). The "w" and "R" in the top left indicate the rotation weight in relation to translation and rotational 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). (B) Top: Illustration of the autoencoder constructed with the translation RFs for determining the direction of corrupted translational optic flow fields (TOFs, see 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 ± the standard deviation. (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 grey lines correspond to the drifted distance when the simulated fish swam randomly.

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. Translational optic flow fields are greatly affected by additional rotational optic flow components (e.g. a shift of the FoC point and the changes of spatial symmetry). This decoder was a variant of the maximum-a-posteriori decoder that used the measured RFs and the neural responses to optic flow in Figure 3E to estimate the translation direction of the presented optic flow combination. The estimation was achieved by minimizing the squared difference

between the recorded responses of the translation-sensitive neurons and the prediction by their RFs (Figure 5A, see Method for details).

The similar decoding error distribution for translational optic flow with or without rotational optic flow interference suggests a robust encoding of translation direction (Figure 5A bottom). Indeed, AMC responses to the pure translational optic flow and the mixed optic flow looked very similar (Movie 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 (Figure S7A,B, see Methods for details). In contrast to the low decoding error for translation direction (median error: 6.6 degrees, the decoding error for translation speed was high (8.56 degree/second or 32.6% of the stimulus speed) when compared to 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 translation optic flow fields with 100 different translation directions at 4 completeness levels (25 %, 50 %, 75 %, 100 %) corresponding to the proportion of the visual field in which optic flow is visible. For each TOF, an autoencoder was constructed with 50 RFs randomly sampled from the total 329 translation RFs (see 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 degrees. It was only 5.7 degrees higher than the one for the 100% completeness level (8.7 degrees, 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 S7E). 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 behaviours, e.g. for self-stabilizing behaviour in water currents (Roeser and Baier, 2003). 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 the identified 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 flow fields (Figure 5C, also see example fish in Movie 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 behaviour.

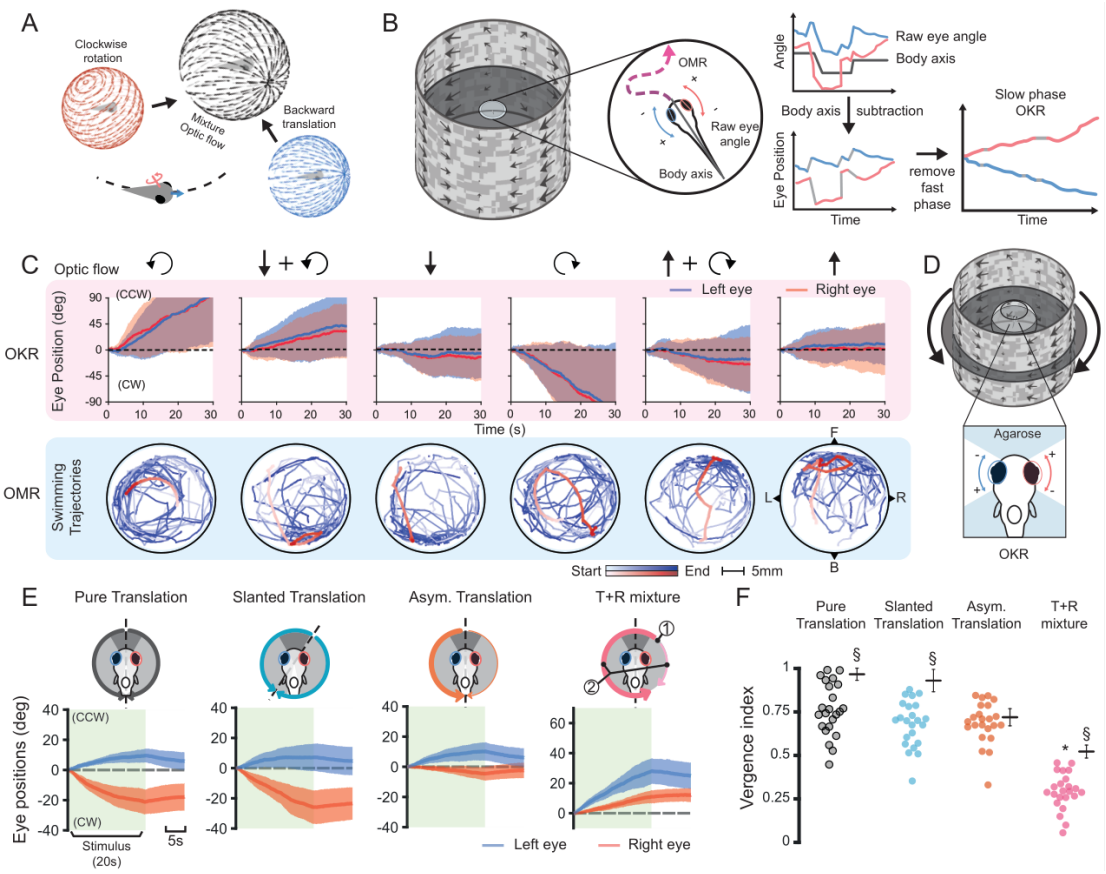

Figure 6 Zebrafish decompose mixed rotation and translation optic flow in optokinetic and optomotor responses (OKR, OMR). (A) Illustration of the mixture optic flow field resulting from the counter-clockwise (CCW) turning of a forward moving fish. (B) Setup for simultaneous OKR (eye movements) and OMR (swimming trajectory) recording for freely swimming fish (see Methods). For OKR (right), saccades were removed to analyse 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: standard deviation across all fish). One example swimming trajectory is coloured in red for each stimulus. “F”, “B”, “L”, “R”: the forward / backward / left / 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 ± standard deviation) across 18 trials recorded in five semi-restrained fish. The OF stimuli are illustrated on top. Dark grey 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 focus of contraction (FoC) 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 ± std. error bars) expected for independently moving eyes responding to the average stimulus directions and speed in each hemi-field (using bootstrapping of local stimulus

statistics, see Methods). 1 and 0 correspond to pure convergent/divergent and pure conjugate movement, respectively. The asterisk indicates significantly lower vergence indices compared to the other three groups. The "§" indicates a significantly lower measured ocular vergence index in comparison to the corresponding local flow statistic. The Ranksum test (α = 0.05) and Bonferroni correction was used to assess significant differences.

Distinct optokinetic and optomotor responses to translational and rotational components in complex optic flow field

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). Now the neural mechanisms for optic flow processing have been revealed. The robust encoding of translation direction (Figure 5A and Movie S2) against the interference of rotational optic flow component strongly suggests translation and rotation information from the optic flow input are segregated in fish 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 prove this, we selected the optomotor responses (OMR) and optokinetic responses (OKR) as two behaviours readouts as they are tightly linked to optic flow perception and represent prominent stabilization behaviours in zebrafish (Kubo et al., 2014; Roeser and Baier, 2003). We recorded the OKR and OMR of 24 freely swimming fish to pure forward/backward translational and CW/CCW rotational optic flow, as well as to combinations of both.

As shown in Figure 6C and S8, fish responded to pure rotation with both OMR and OKR, as indicated by their circular swimming trajectories and conjugate eye movements. For the pure translation optic flow, fish swam with the optic flow direction, while clear conjugate or vergence eye movements were absent. 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 (Figure 6 and S8).

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 translational optic flows. For optic flow combinations, both semi-restrained and freely swimming moving fish showed conjugate eye movements without significant vergence components (Figure 6E). Together, these behavioural evidences 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: are the FoC 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 translational optic flow. In the “slanted translation” variant, the translation direction of the pure translational optic flow field was shifted so its FoC was at the same position as the one in the mixture of translation- and rotation-induced optic flow field. 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 vs. vergence eye movements using a vergence index (Figure 6F, 1 = pure vergence movement; 0 = pure conjugate movement; see method for details). The vergence index for the translation optic flow combined with a rotation component was significantly shifted towards conjugate eye movements in comparison to 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 to 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 fine-grained receptive field structure of motion-sensitive neurons in the optic tectum and pretectal area, we revealed that many neurons, especially pretectal ones, perform large-scale motion integration with their bimodal RFs. Most of these RFs belong to a novel RF class, complex RFs, and play a pivotal role in encoding translational and rotational optic flow. By revealing the algorithms implemented in these complex RFs for optic flow encoding, we successfully demonstrate how the encoding properties, such as translation/ rotation selectivity may emerge from these complex RFs. We used this information to uncover a topographic map of preferred translation directions in the AMC region of the pretectal area, and most of the motion-sensitive neurons in this map were found to jointly encode both the speed and the direction of translational optic flow.

Furthermore, we demonstrate that the information encoded by the translation-sensitive neurons with their complex RFs is sufficient for determining the direction and speed of the translational self-motion component. Moreover, the encoding is robust against corruption of the optic flow field and against interference from additional rotational components. Finally, optokinetic and optomotor behavioural responses clearly showed that zebrafish decompose the translation and rotation components from the combinatory optic flow.

Complex RFs implement the optic flow detection algorithm

The perception of self-motion requires large scale integration of local motion information. The corresponding neural correlates for this integration have been discovered in many different regions in the visual system from insects to primate (Eckmeier et al., 2013; Egelhaaf, 2002; Horseman et al., 2011; Huston and Krapp, 2008; Joesch et al., 2008; Stone et al., 2017; Sunkara et al., 2016). It has been proposed for a long time that there is a tight link between the RFs and the optic flow tuning properties of corresponding neurons. However due to the throughput limit of RF estimation methods, this link has only been demonstrated in insects with a small number of neurons involved (Egelhaaf, 2002; Franz and Krapp, 2000). 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 3926 motion RFs reported here, hinged on the availability of a fast, multiplexed RF estimation technique developed for calcium imaging (Zhang and Arrenberg, 2019). The RF structures of the global motion sensitive neurons revealed in this study showed remarkable similarity to translational and rotational optic flow fields (Figure 3C). This observation is consistent with the previous studies lobula plate tangential cells (LPTC) in flies for which a “matched filter” model has been proposed (Franz and Krapp, 2000; Krapp and Hengstenberg, 1996; Krapp et al., 1998). In short, this model suggests the LPTC neurons may filter impinging optic flow fields with their RFs, resulting in a stronger response for an optic flow field matching the RF structure (Franz and Krapp, 2000). The calcium responses of our complex RF neurons, especially the ones with translation RFs, are well described by this model

(Figure 3F, Figure S4). Furthermore, we establish how the translation/rotation preference may emerge from complex RF structure as a result of this matched filter and show that the level of remaining unexplained response modulation is small for most of the complex RF neurons.

In contrast, the responses of the simple RF neurons frequently differed from the predictions of the matched-filter model. For a fraction of neurons, the responses to optic flow mixtures were even opposite to the prediction based on the measured RF (correlation <0 in Figure 3F, Figure S4). This finding is quite unexpected because their RFs, as the name suggests, should be rather simple. The poor prediction for simple RFs suggests that our RF estimates were missing important functional aspects for these neurons. One possibility is that many of these neurons had suppressive RF components remaining undetected in our RF estimation. In neurons with low baseline activity, inhibition (or in general activity suppression) is notoriously hard to demonstrate via calcium imaging (Zhang and Arrenberg, 2019). A previous study (Wang et al., 2020) showed that about 80% of the motion-sensitive neurons in the optic tectum have suppressive surround structures, i.e. they were suppressed when the motion stimulus covered too large a region of visual space. Therefore, it is very likely that many of our simple RF neurons had additional suppressive components in their RFs. Similarly in the salamander retina, some local motion sensitive retinal ganglion cells (RGC) are also suppressed by large scale motion (Kühn and Gollisch, 2016). 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 arranged according to preferred translation directions in a single topographic map that spans both pretectal hemispheres (Figure 4B & S6E, Movie S1 & S2). Previous reports already revealed differential directional processing across pretectal locations and related these direction-selective neurons to the triggering of optomotor responses (Kubo et al., 2014; Naumann et al., 2016; Wang et al., 2019, 2020; Yildizoglu et al., 2020).

These results agree with our findings despite major differences in employed visual stimulus setups. In the studies by Naumann et al. (2016) and Yildizoglu et al. (2020), motion stimuli were projected on a flat screen from below, which can be interpreted as the bottom part of a translational optic flow field. For example, the nasal-temporal-ward motion covering the bottom visual field can be interpreted as the bottom part of a translational optic flow field 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 (Naumann et al., 2016, Figure 3A) in terms of the preferred direction 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° 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 behaviour. Larva zebrafish can adjust their swimming behaviour to match the speed of wide-field visual motion stimuli (Portugues et al., 2015; Severi et al., 2014). 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 translational optic flow is usually conjugated with the depth and location of stimuli in the visual field (Koenderink and van Doorn, 1987). Whether and how do translation-sensitive neurons in the AMC regions untangle this information for determining self-motion speed requires further investigation.

Robust and unambiguous coding of translation direction in translation RFs

The estimation of the translation direction requires the integration of motion from different parts of the visual field. In principle, this information can be encoded either by the complex translation RFs as shown in this study, or a population of neurons tuned to more simple and basic motion features, such as the direction-selective retinal ganglion cells (DSRGCs) in mouse retina described in a previous study (Sabbah et al., 2017). However, with the latter encoding strategy, the translation direction cannot be extracted as a standalone feature because of the encoding ambiguity due to the absence of translation selectivity. In other words, without additional decoders, it is not possible to discriminate translational optic flow from other types of optic flows from the population code. Carlson and Kawasaki (2006) showed – for a different sensory modality - how such an ambiguous representation of stimuli will eventually result in the confusion in behavioural responses to distinct features. And this is not observed in zebrafish responses to the mixture of translational and rotational optic flow (Figure 6). Therefore, we hypothesize that our identified population of pretectal neurons with complex RFs directly extract and convey the translation direction to the motor system. These neurons and thus forms a visual encoding pivot before information is passed on to premotor areas.

Furthermore, in Figure 5 we showed the population code of translational optic flow is robust against corrupted of optic flow field. This explained why motion in relatively small parts of the visual field can already trigger strong OKR and OMR (Dehmelt et al., 2020; Kist and Portugues, 2019; Wang et al., 2020), which is a result that seems puzzling at first glance when considering the optomotor response as a “whole field” stabilization behaviour.

Represented optic flow axes

We noticed there are some invariant properties across RFs in the optic tectum and pretectum: I) the preferred directions 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 Fig. 2E and S2A (Hunter et al., 2013; Lowe et al., 2013; Wang et al., 2019); II) bimodal vertical neurons together encode translation and rotation, where most bimodal horizontal neurons encode translation; III) the modes of bimodal translation RFs are mostly distributed symmetrically in relation to the preferred translation direction which does not coincide with the animal's longitudinal body axis as shown in Figure 2C and 4E; And IV) most bimodal RF centres are located 45° to 90° degrees away from the translation directions (Figure 4D), where motion speed and direction coherence is maximal ("hot zone"). These properties together constrain the preferred translation directions of the bimodal RFs (cf. Figure 4A). Few RFs are tuned to oblique translation directions (combinations of horizontal and lift/weight translations) as these optic flows contain oblique local motion in the "hot zone" of the RFs. It is unclear at this point, what the specific advantages of this layout are. 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 preferred directions. 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 (Fig. 5H). However, spontaneous swimming as well as prey capture behaviour includes frequent locomotor deviations from the horizontal plane (Bolton et al., 2019; Ehrlich and Schoppik, 2017) 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 could be integrated in downstream neurons to signal the presence of rotation stimuli. 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

In the behavioural sequence of prey capture in zebrafish characterized by a recent study, three motion components are superimposed in the retinal projection of the optic flow field: the translation component from the forward swimming behaviour, the rotation component from turning behaviour, and the object motion component from the prey movement (Bolton et al., 2019). The ability of the animal to hunt and orient in the environment suggests that decomposition of these motion components must occur, but how and to what extent flow decomposition occurs in vertebrate animals had been unclear.

While evidence for separate processing of translation and rotation information in the corresponding neural correlate has been presented before (Kubo et al., 2014; Sunkara et al., 2016; Wang et al., 2019; Wylie, 2013), our study shows that translation/rotation sensitive neurons are not exclusively responsive to pure

translational/rotational optic flow, 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 and Movie S2), which indicates that the zebrafish pretectum decomposes translation and rotation components and is biased towards the identification of translation components. In contrast to translation RFs, the responses of rotation-sensitive neurons to mixed optic flow stimuli are often strongly modulated by the non-preferred (translation) components and poorly explained by their RF structure (Figure S4). In a previous study, two types of translation-sensitive neurons have been found in primate parietal cortex: neurons only tuned to translation components of a given optic flow fields, and neurons modulated by both translation and rotation components (Sunkara et al., 2016). The latter are suggested to provide a joint representation of translation direction and rotation velocity which may contribute to the decomposition of translation and rotation components in complex optic flow fields. 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, and what role they do play in optic flow decomposition.

The translation and rotation filters implemented in pretectal neurons provide the substrate for the decomposition of global motion components in visually guided behavior. The OMR and OKR are two prominently studied visually guided behaviours, and their correspondence to translational and rotational self-motion, respectively, have been investigate separately before (Kist and Portugues, 2019; Maaswinkel and Li, 2003; Marques et al., 2018; Qian et al., 2005; Rinner et al., 2005). However, it was unknown whether the translation- and rotation- induced optic flow components, can guide the OMR and OKR in separate ways, or in other words, do zebrafish "decompose" these global motion components in their OKR and OMR behaviour? Here we answered this question by recording OKR and OMR simultaneously in freely swimming larvae. Fish clearly processed the translation- and rotation-induced components in the presented optic flow stimulus separately, as they were swimming toward the corresponding translation direction implemented in the optic flow field while performing OKR in the direction of rotation. Apparently the OMR is predominantly modulated by the translation-induced components (Figure 6C), while the OKR is more sensitive to the rotation-induced components of the optic flow stimuli (Figure 6C, E and F). The local velocity vector statistics of the rotation-induced optic flow component, but not the translation-induced optic flow component does not vary by horizontal fish body orientation. This difference is also reflected in the OKR of freely swimming zebrafish: for the pure translational optic flow we observed vergence OKR in head mounted fish but not in freely swimming fish, while the conjugate OKR to the optic flow stimuli with rotational component can be observed in both cases. This result also implies the discrimination of optic flow with and without rotational component does not rely on local features at certain positions of the visual field which will vary during head direction changes. Instead, the isolated local features resulting from rotational component interference were unable to trigger conjugate eye movements (Figure 6E&F), suggesting the integration of local features but not a specific local feature is essential for the optic flow processing in zebrafish. It will be interesting to determine the full neural circuit for this

decomposition in the future, including the connectivity of translation-sensitive neurons in AMC (Kramer et al., 2019)

In summary, we have systematically characterized the receptive field 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 behaviours.

689 # Materials and methods

690

691 ## Ethical approval

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

695

696 ## Animal preparation

697 Zebrafish (Danio rerio) were maintained on a 14/10h light–dark cycle at 28°C. The
698 transgenic zebrafish line *Tg(elavl3:nls-GCaMP6s)mpn400* was used for the *in vivo*
699 two-photon calcium imaging (Dal Maschio et al., 2017). The top 10 - 20% of larval
700 zebrafish with high calcium indicator expression levels and strong optokinetic
701 responses (OKR) to moving gratings were selected from all hatched fish at 4-day
702 post-fertilization (dpf) for *in vivo* calcium imaging experiments.

703 5-7 dpf zebrafish larvae carrying mutations in the mitfa gene (nacre) was used for
704 OKR and optomotor responses (OMR) recordings. Again, the top 10-20% larval
705 zebrafish sensitive to visual motion stimuli indicated by their strong OKR to moving
706 gratings were selected at 4 dpf for these behavioural experiments.

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

709

710 ## Visual stimulation setup

711 All visual stimuli used in this study were displayed on a custom-built cylindrical green
712 LED arena (160 mm height, inner diameter = 198 mm) covering 336°-by-80° of the
713 visual field (-168° to 168° in azimuth; -40° to 40° in elevation). The 2 halves of LED
714 arena consisted of 224 square LED tiles (8 rows × 14 columns × 2 arena halves); each
715 tile had 8×8 evenly distributed LEDs emitting at 570nm (Kingbright TA08-81CGKWA).
716 Each LED covered ˜1.5 visual degrees on average, note the coverage on vertical
717 direction varied slightly with the latitude of LEDs. All stimuli displayed in the arena
718 had 100% contrast as the LEDs in the stimuli were either on or off. All LED tiles were
719 covered with a diffusion filter foil (LEE no. 252, article 595-1780-2520) and a high-
720 pass filter foil (LEE no. 779, article 595-1700-7790, castinfo.de, Hagen, Germany) for
721 light homogenization and light interference reduction in calcium imaging
722 respectively.

723 The custom software developed in previous studies (Joesch et al., 2008; Reiser and
724 Dickinson, 2008) was used for stimulus uploading and display control. To minimize
725 the interference of LED light with the calcium imaging, the LEDs were flickering in a
726 duty cycle to be switched on only during the fly-back time of the scanning mirrors in-
727 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 (Zhang and Arrenberg, 2019). 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 (Zhang & Arrenberg, 2019). 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 (Rinner et al., 2005).

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:

$$\vec{V}_{rot} = -\vec{P} \times \vec{R},$$

$$and\ \vec{V}_{trans} = \vec{T} - (\vec{P} \cdot \vec{T})\vec{P}$$

$\vec{P}$ 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 <×> and <·> 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:

$$\begin{aligned} \vec{V}_{corrected} &= \mathbf{D}\left(\vec{V}_{raw} - \left(\vec{V}_{raw} \cdot \vec{N}\right)\vec{N}\right), \\ \text{where } \vec{N} &= \frac{(x, y, 0)}{|(x, y, 0)|}, \\ \text{and } \mathbf{D_i} &= \frac{r}{\cos(\theta)} = \frac{r}{\sqrt{1 - z^2}} \end{aligned}$$

, 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 (circled region in the illustrations of Figure S6A).

In vivo two-photon calcium imaging and data pre-processing

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 S9) on the experiment day. The triangle stage with the mounted fish was later placed in the centres of a spherical glass container filled with transparent E3 medium. Then the water container was fixed on a metal holder placed in the centres of the LED arena. The body position and orientation of the mounted fish was adjusted so it was centred in the LED arena with the dorsal side facing up (roll

correction) and the noise pointing straight to the front (0° azimuth and elevation) without any tilting.

The calcium activities in neuron soma 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 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).

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 artefacts along the recording plane (XY plane) were corrected, and all visible neuron soma in the time-averaged image of the corrected recording video were automatically selected as region-of-interests (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 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 (Zhang and Arrenberg, 2019). Briefly, the neural responses to the CMN stimulus were recorded with *in vivo* calcium imaging. As the abrupt, high amplitude increases of somatic 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:

$$ETA = \frac{1}{n_{event}} \Sigma_{i=1}^{I} e_i B(\vec{V_i})$$
$$n_{event} = \Sigma_i^I e_i$$

The e_i was the event state (0 or 1) for the ith 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 ith 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 (see Zhang & Arrenberg, 2019 for details). For instance, for a vector $\vec{v}$ in $\vec{V_i}$ whose phase angle is in the range of the kth 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 (see Zhang and Arrenberg, 2019 for detailed explanations). For the sake of clarity and convenience, the identified 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 centres and preferred directions of RF modes were calculated as the averaged location and preferred direction of the components in each mode.

Fitting translational and rotational 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:

$$RSS_{angle} = \Sigma_{i=1}^{I} \angle(\vec{E_i}\vec{F_i})^2$$
$$\angle(\vec{E_i}\vec{F_i}) = \arccos\left(\frac{\vec{E_i} \cdot \vec{F_i}}{|\vec{E_i}||\vec{F_i}|}\right)$$

And the fitting problem can be solved by searching the translation/rotation axis that minimized the RSS_{angle}. 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:

$$\text{Similarity} = \frac{1}{I} \Sigma_{i=1}^{I} \frac{\vec{E_i} \cdot \vec{F_i}}{|\vec{E_i}||\vec{F_i}|}$$

In addition, the symmetry of bimodal translation RFs in relation to the preferred translation directions in Figure 4E was calculated as:

$$\text{Symmetry} = 1 - abs\left(\frac{2\angle(\vec{\Delta}_{center}, \vec{T})}{\pi} - 1\right),$$

$$where\ \vec{\Delta}_{center} = \mu_{left} - \mu_{right}$$

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 ($\angle(\vec{\Delta}_{center}, \vec{T}) = \frac{\pi}{2}$).

Optic flow preference of motion-sensitive neurons

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 to the translational and rotational optic flow in 6 orthogonal directions and the 24 combinations of these translational and rotational flow displayed 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 coloured in white in Figure 3E. 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.

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 and strong responses to optic flow stimuli ($max(R) > 3$) were selected. The RF-predicted responses ($R_{prediction}$) to the presented optic flow stimuli ($\overrightarrow{GF}$) were calculated as follows:

$$R_{\text{prediction}} = max\left(\frac{1}{N}\Sigma_{n=1}^{N}\frac{\overrightarrow{RF_n}\cdot\overrightarrow{GF_n}}{\left|\overrightarrow{RF_n}\right|\left|\overrightarrow{GF_n}\right|}, 0\right)$$

N is the number of significant components in the RF estimations, $\overrightarrow{RF_n}$ and $\overrightarrow{GF_n}$ are the vectors of the nth 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.

Direction and speed tuning of translation sensitive neurons

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 S6A) 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 (DeAngelis and Uka, 2003):

$$\text{Discrimination Index} = \frac{R_{max} - R_{min}}{R_{max} - R_{min} + 2\sqrt{SSE/(N-M)}},$$

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.

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:

$$\text{von Mises-log Gaussian distribution} = \overbrace{\frac{1}{v\sigma\sqrt{2\pi}}\exp-\frac{\ln(v-\mu_s)^2}{2\sigma^2}}^{\text{Log Gaussian term}}\overbrace{\frac{e^{\kappa cos(\theta-\mu_d)}}{2\pi I_0(\kappa)}}^{\text{von Mises term}}$$

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 and autoencoder for translational optic flow 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 (Abbott and Dayan, 1999), 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 by least square fitting which minimizes the sum squared residue between the calcium responses R of the recorded n translation-sensitive neurons, and the responses predicted by their response functions ($F(T)$):

$$\hat{T} = argmin\left(\sum_{i}^{n}(F_i(T)-R_i)^2\right)$$

The responses functions were approximated with either the RF-based response function (Figure 5A bottom, see the section "Optic flow preference of motion-sensitive neurons" for details) or the joint distribution model fitted to the direction-speed tuning data (Figure S6B-D, S6F, S7A-B). 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 S6B, 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 jackknife test ("leave-ten-out") to avoid this circular analysis: in the decoder for Figure S7A-B, 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 Figure 5B and S7E) with *n* translation RFs into a one-dimensional array of neural responses as:

$$R_j = F_j(S) + \varepsilon\,,$$

$$where\ F(S) = max\left(\frac{1}{I}\Sigma_{i=1}^{I}\frac{\overrightarrow{RF_i}\cdot\overrightarrow{GF_i}}{\left|\overrightarrow{RF_i}\right|\left|\overrightarrow{GF_i}\right|}, 0\right)$$

R_j and $F_j(S)$ are the simulated recorded response and the RF-based response function of the jth simulated neurons, 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 was tested with different encoding signal-to-noise ratio (SNR = +∞ (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.

We also evaluated the encoding quality when the number of translation RFs involved *n* = 10, 20, 50, 100 and 329. Based on the results in Figure S7C-D, the autoencoder used in Figure 5B and S7E 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 S10). The rest of the optic flow field were occluded or replaced with random object motion (Figure 5B and S7E). 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 $\vec{O}$ indicates the object motion in the 3D environment, the motion vector $\vec{V}$ in a region of the visual field indicated by a unit vector $\vec{P}$ was calculated as:

$$\vec{V} = \vec{P} - (\vec{P} \cdot \vec{O})\vec{O}$$

Simultaneous recording of eye movement and body position in freely swimming zebrafish

The fish at 5-6 dpf with strong behavioural 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 centres 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 artefacts. This reservoir was placed in the centres 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, therefore we do not expect any strong disturbance from internal reflection to animal behavior as discussed in a previous paper (Wang et al., 2021). 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 were 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 centres 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 which is perpendicular to the line linking the centres of two eyes and pointing into the coarse body direction determined previously. The orientation of the two eyes were 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 and detection 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 centres of the cylindrical LED arena. The body position and orientation of the mounted fish was adjusted later, so it was in the centres 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 and behavior data analysis for optic flow decomposition experiments

As illustrated in Figure 6C, 8 different optic flow stimuli were used in the global motion decomposition experiments in freely swimming fish, 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:

$$\boldsymbol{D} = \sqrt{1 + \left(\frac{w_{rot}}{w_{trans}}\right)^2 - 2\,\frac{w_{rot}}{w_{trans}}\cos\left(\angle\vec{P}_{azi}\right)}$$

$\angle\vec{P}_{azi}$ is the azimuth angle of the unit vector $\vec{P}$ 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.

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.

1141

1142 **Retrograde labelling of the nucleus of the medial longitudinal fasciculus (nMLF)**

1143 The retrograde labelling of nMLF in nls-GCaMP6s zebrafish larvae was performed as described
1144 previously (Thiele et al., 2014). Briefly, a 50% Texas Red dextran (10000MW, Thermo Fisher)
1145 in Evans solution was injected into the spinal cord of 4 dpf. larvae embedded in 1.7% agarose
1146 and anesthetized with tricaine. The injected fish was imaged with the two-photon microscope
1147 at 5 dpf.

1148

Supplementary Figures

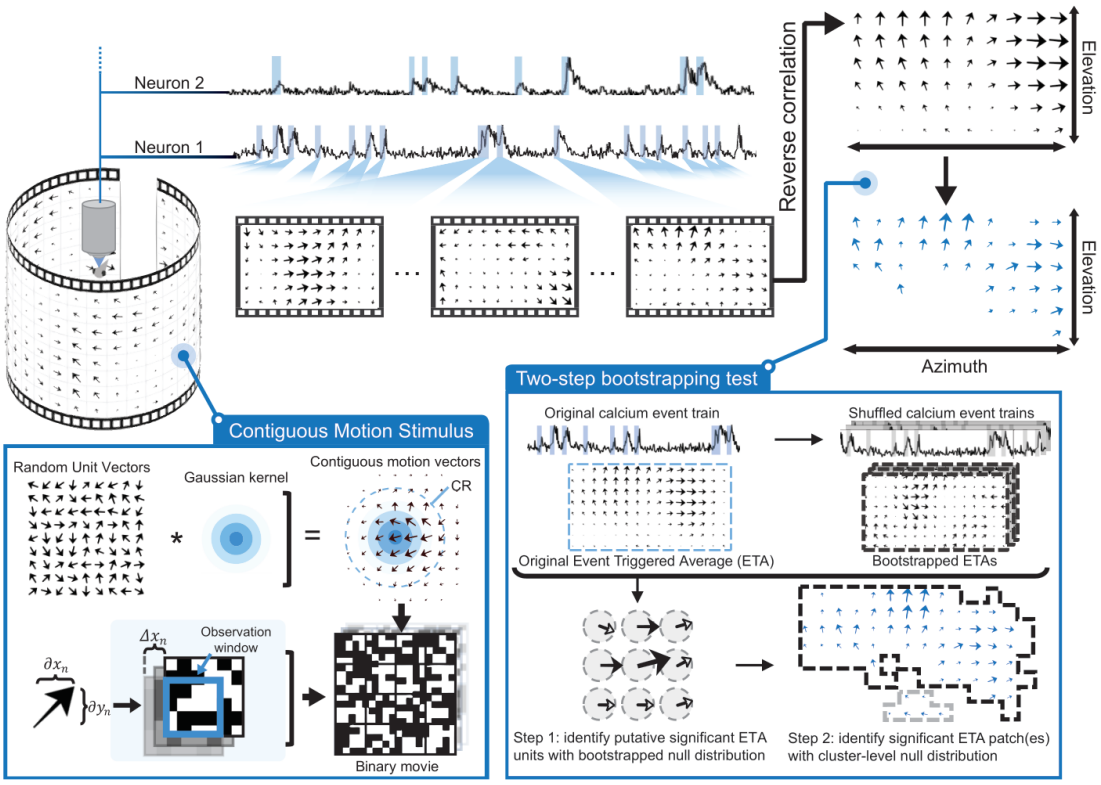

Figure S1. A graphical abstract for the receptive field estimation method using contiguous motion stimulus and two-step bootstrapping test (adapted from Zhang and Arrenberg, 2019, see Methods for details).

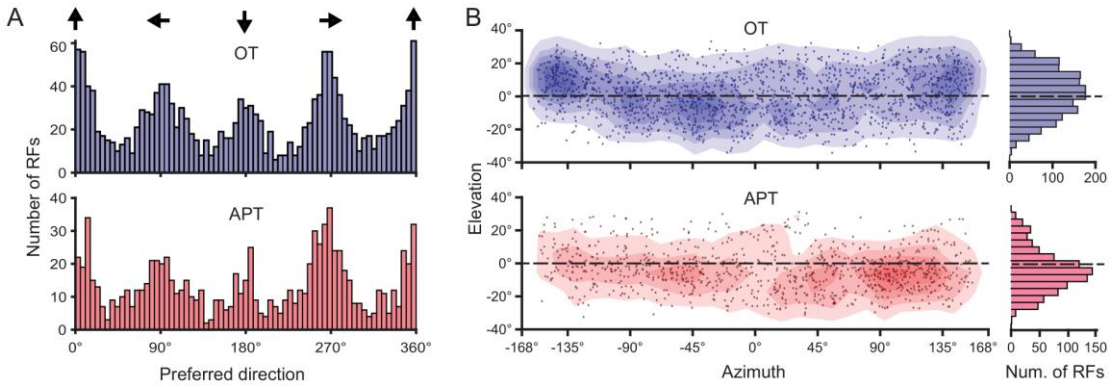

Figure S2. The distribution of averaged preferred directions (A) and RF centre locations (B) in the visual fields for the neurons with unimodal RFs in optic tectum (OT) and area pretectalis (APT). The black arrows in the top row of A indicated the preferred motion direction in 2D. Each dot in (B) indicates 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.

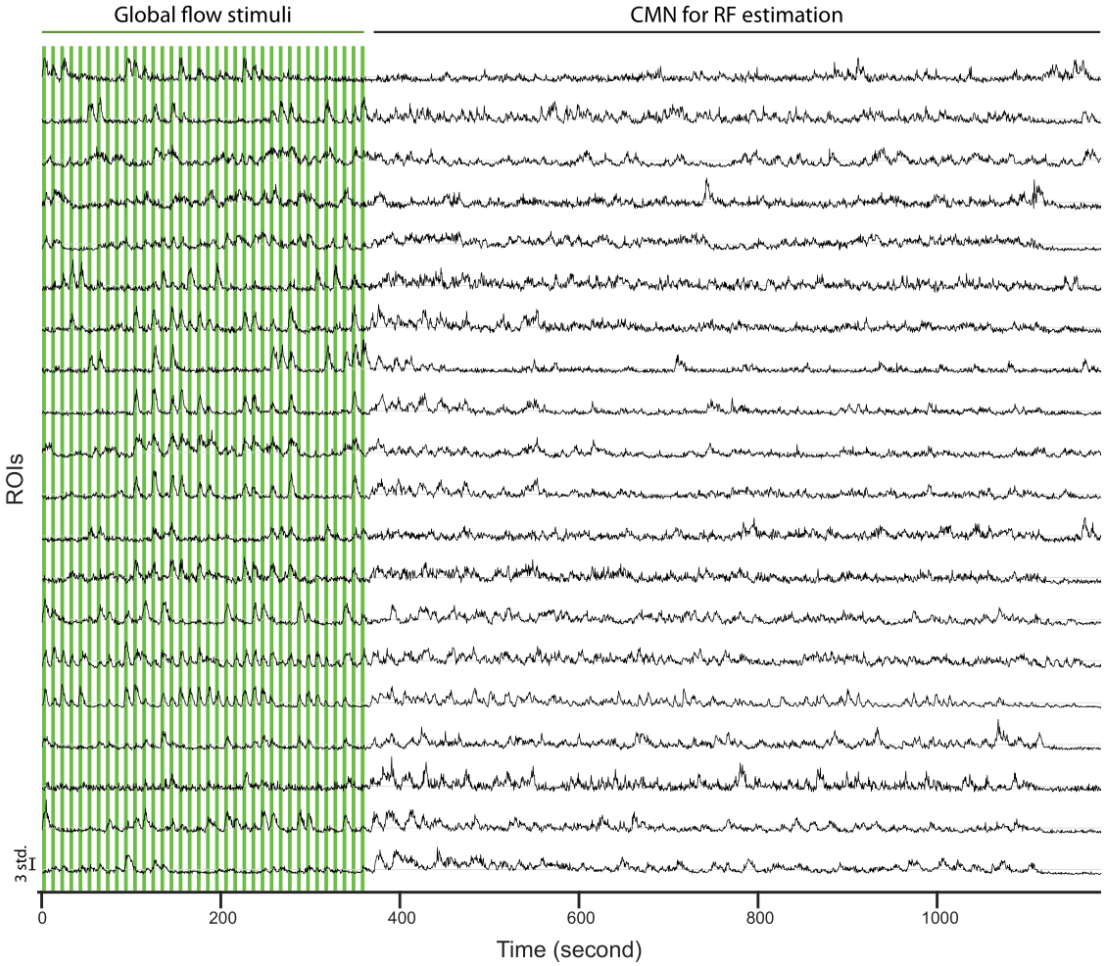

Figure S3. Examples of z-score calcium traces recorded in the experiment for Figure 3E-G and Figure S3. Each green shaded area indicates an optic flow stimulus phase (4 seconds duration + 6 seconds pause interval).

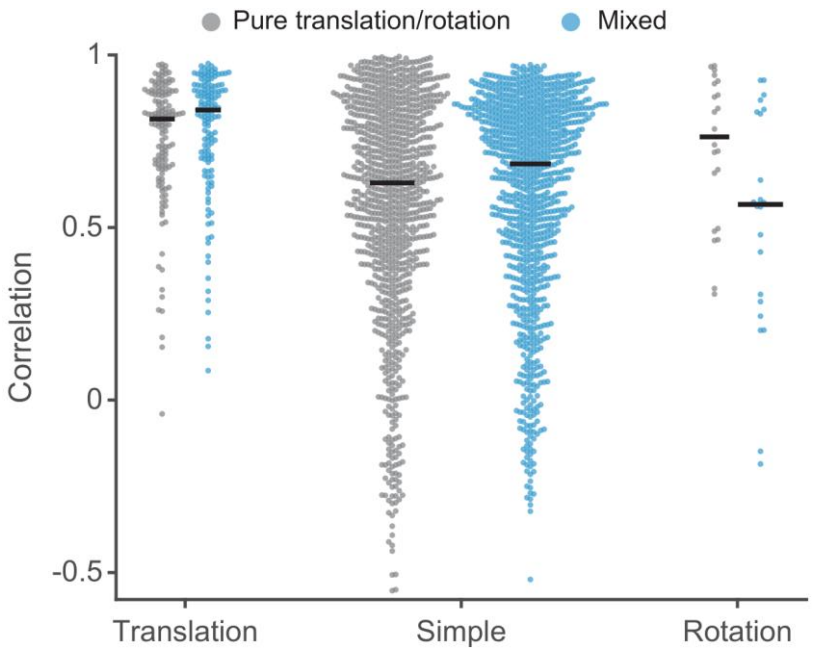

Figure S4. The bee-swarm plots of the correlation between the predicted and real neural response intensity to pure translational/rotational or mixed optic flows used in Figure 3E. The neurons were grouped by their RF types indicated in the x axis.

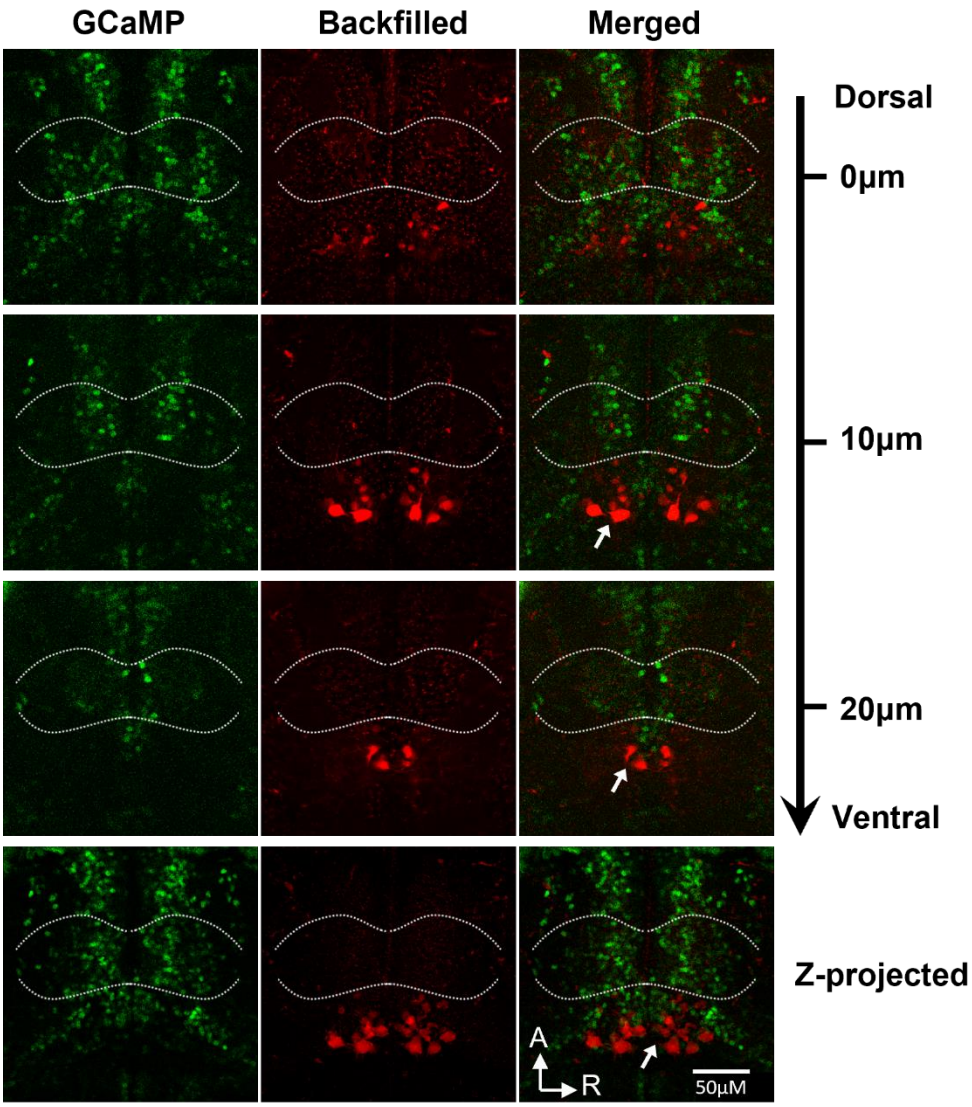

Figure S5. Dorsal view of the pretectal area in 3 different depths. The green channel (GCaMP) displays all visible neural somata (two-photon calcium image) in the corresponding area, the white dashed curves outline the approximate location of the AMC pretectal clusters. The red channel shows the neural structures retrogradely labelled by spinal cord injections. 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.

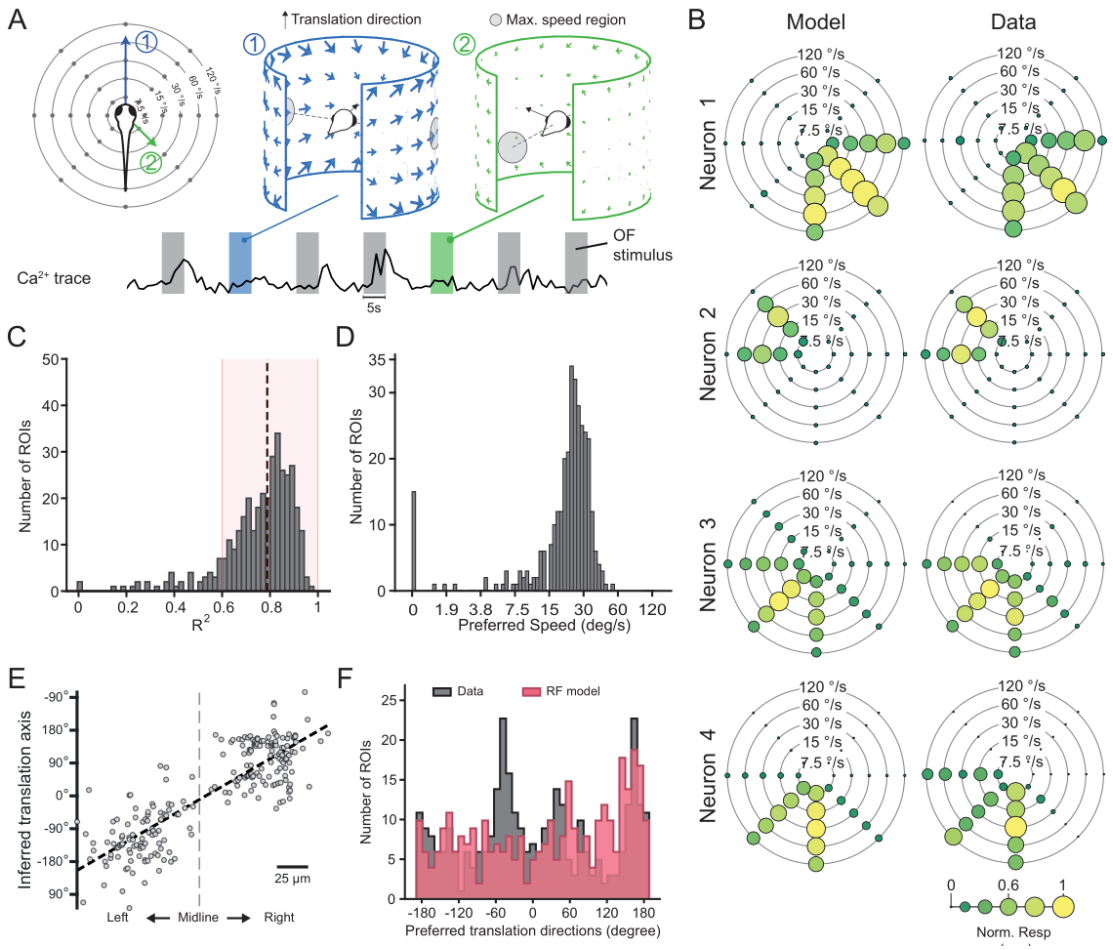

Figure S6. Joint encoding of translational optic flow direction and speed in the pretectum. (A) Left: Animals were present with stimuli consisting of one of eight different TOF directions and five logarithmically spaced maximum velocities. Right: Two examples TOF stimuli are shown (① and ②). The region 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 indicated 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) Linear regression (black dashed line, R^2 = 0.714) 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. (F) 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")

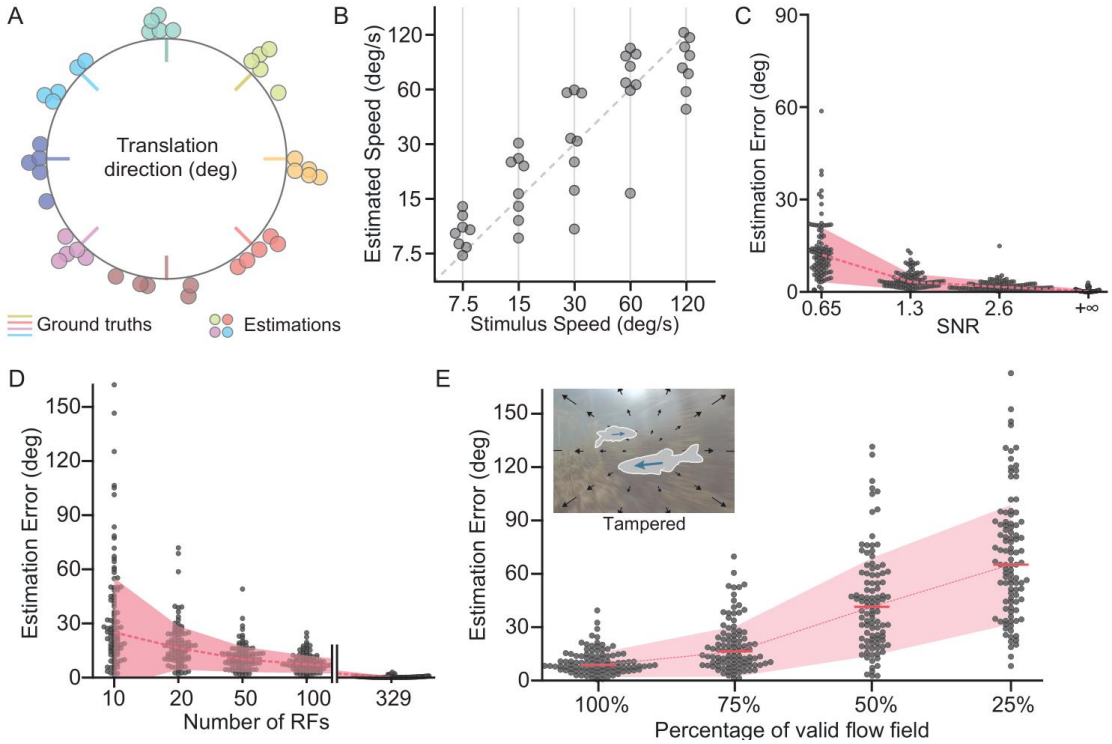

Figure S7. (A) The translation direction estimated with the data presented in Figure 4E-F plotted in bee swarm style. Colored bars indicate ground truth horizontal translation directions and circles the estimations. (B) The translation speed estimated with the data presented in Figure 4E-H in bee swarm 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.

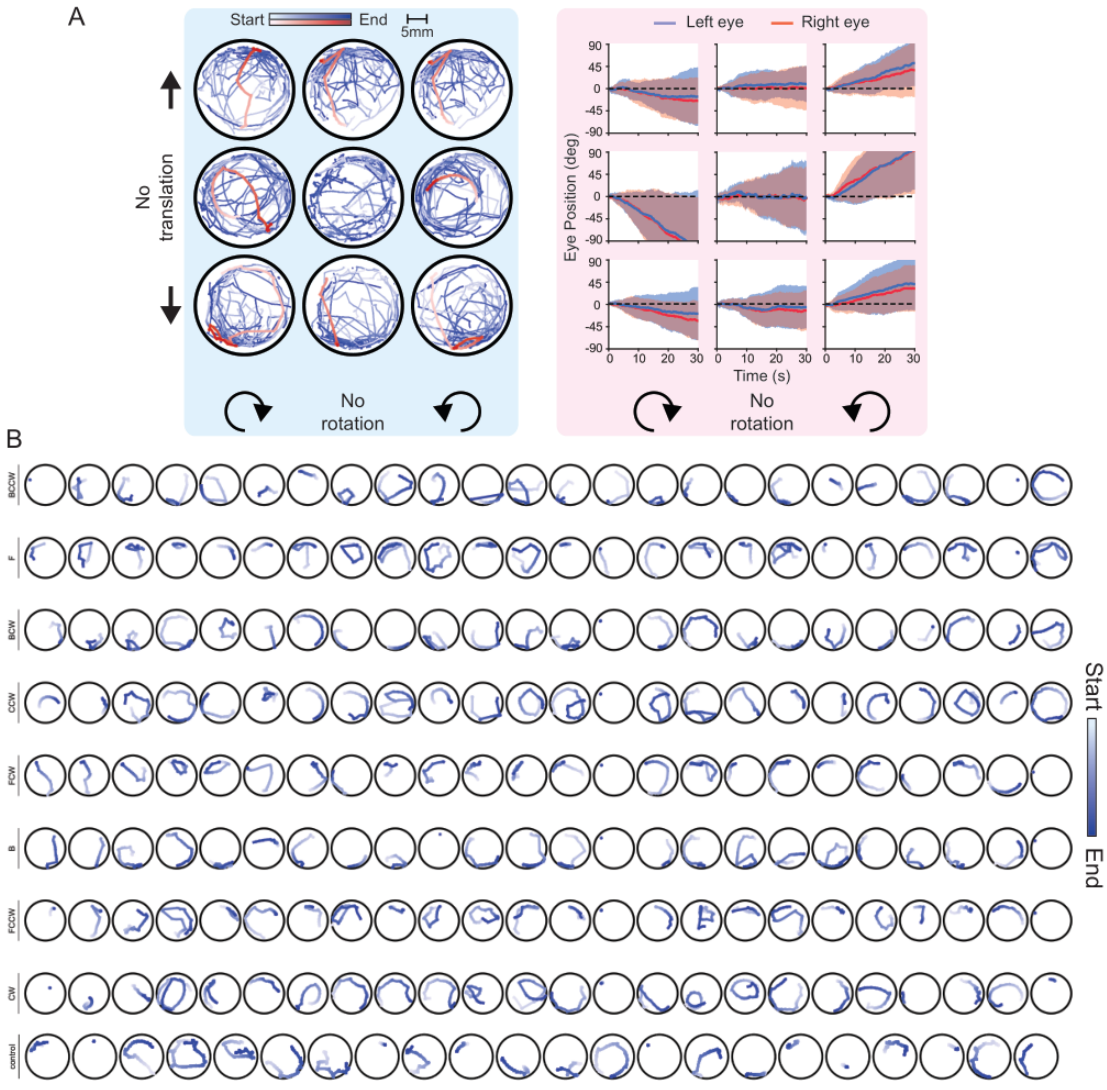

Figure S8. (A) The superimposed swimming trajectories (left) during all 9 optic flow stimuli phases (indicated by the arrows at the sides) and the corresponding eye position traces (mean ± standard deviation). (B) The individual swimming trajectories for each optic flow stimulus (rows) and fish (column). From top to bottom: BCCW (backward translation + counter-clockwise rotation), F (forward translation), BCW (backward translation + clockwise rotation), FCW (counter-clockwise rotation), B (backward translation), FCCW (forward translation + counter-clockwise rotation), CW (clockwise rotation) and control.

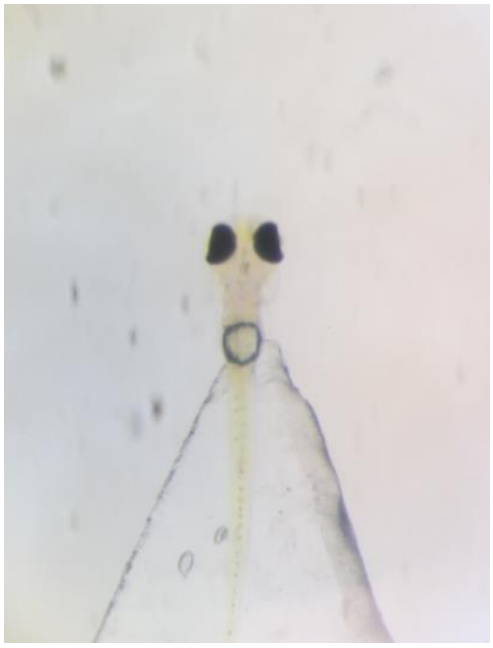

Figure S9. An image of an agarose-mounted 5 dpf zebrafish larva on the tip of a plastic triangle stage as used in this study.

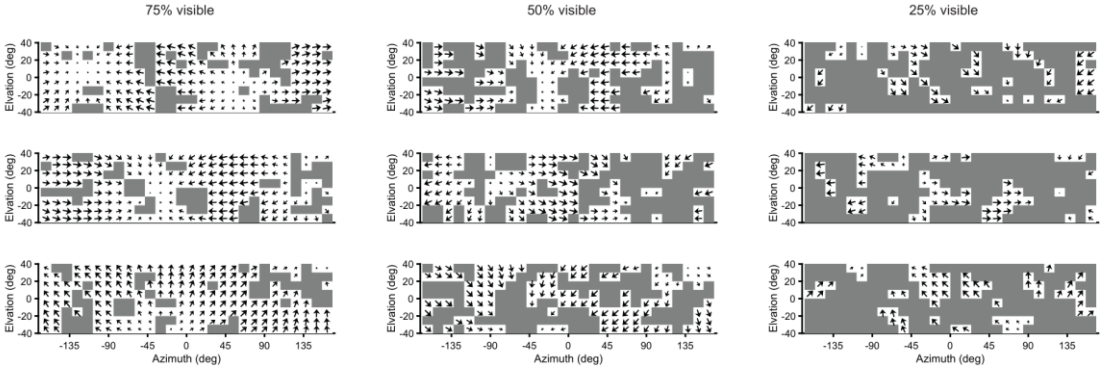

Figure S10. Examples of incomplete or tampered translational optic flow fields. Each gray block indicates a region in the visual field that contains no or wrong motion cue. In the tampered visual field, each set of spatially connected gray blocks displayed the 2D projected motion resulting from the 3D movement of a single object covering this region.

MovieS1.avi

Movie S1. The calcium activities of neurons in the pretectal area in response to pure translational optic flow, whose instantaneous direction is indicated by the spinning arrow.

MovieS2.avi

Movie S2. The calcium activities of the same neurons in Movie S1 in response to mixed optic flow (translation + rotation). The instantaneous direction of the translation component is indicated by the spinning arrow.

MovieS3.avi

Movie S3. An example animation of a simulated fish stabilized its position in constant simulated "water current" by swimming against the estimated translational optic flow directions.

MovieS4.avi

Movie S4. An example clip of the OF stimulus (2D projected) used in Figure 4E.

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.

Declaration of interests

All authors declare that they have no competing interests.

References

Abbott, L.F., and Dayan, P. (1999). The Effect of Correlated Variability on the Accuracy of a Population Code. Neural Computation *11*, 91–101.

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*.

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.

Britten, K.H. (2008). Mechanisms of self-motion perception. Annu Rev Neurosci *31*, 389–410.

Carlson, B.A., and Kawasaki, M. (2006). Ambiguous Encoding of Stimuli by Primary Sensory Afferents Causes a Lack of Independence in the Perception of Multiple Stimulus Attributes. J. Neurosci. *26*, 9173–9183.

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. (Berlin, Heidelberg: Springer), pp. 143–168.

DeAngelis, G.C., and Uka, T. (2003). Coding of Horizontal Disparity and Velocity by MT Neurons in the Alert Macaque. J Neurophysiol *89*, 18.

Dehmelt, F.A., Meier, R., Hinz, J., Yoshimatsu, T., Simacek, C.A., Wang, K., Baden, T., and Arrenberg, A.B. (2020). Gaze stabilisation behaviour is anisotropic across visual field locations in zebrafish. BioRxiv 754408.

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. Journal of Neurophysiology *65*, 1329–1345.

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.

Eckmeier, D., Kern, R., Egelhaaf, M., and Bischof, H.-J. (2013). Encoding of naturalistic optic flow by motion sensitive neurons of nucleus rotundus in the zebra finch (Taeniopygia guttata). Front. Integr. Neurosci. *7*.

Egelhaaf, M. (2002). Neural encoding of behaviourally relevant visual-motion information in the fly. Trends in Neurosciences *25*, 96–102.

Ehrlich, D.E., and Schoppik, D. (2017). A novel mechanism for volitional locomotion in larval zebrafish. BioRxiv 189191.

Festl, F., Recktenwald, F., Yuan, C., and Mallot, H.A. (2012). Detection of linear ego-acceleration from optic flow. Journal of Vision *12*.

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.

Gibson, J.J. (1954). The visual perception of objective motion and subjective movement. Psychological Review *61*, 304–314.

Graham, D.J., and Wylie, D.R. Zebrin-Immunopositive and -Immunonegative Stripe Pairs Represent Functional Units in the Pigeon Vestibulocerebellum. 11.

Horseman, B.G., Macauley, M.W.S., and Barnes, W.J.P. (2011). Neuronal processing of translational optic flow in the visual system of the shore crab Carcinus maenas. Journal of Experimental Biology *214*, 1586–1598.

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. Journal of Neuroscience *33*, 13940–13945.

Huston, S.J., and Krapp, H.G. (2008). Visuomotor Transformation in the Fly Gaze Stabilization System. PLOS Biology *6*, e173.

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. Current Biology *18*, 368–374.

Jr, W.H.W., Kay, B.A., Zosh, W.D., Duchon, A.P., and Sahuc, S. (2001). Optic flow is used to control human walking. Nature Neuroscience *4*, 213–216.

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

Koenderink, J.J., and van Doorn, A.J. (1987). Facts on optic flow. Biol. Cybern. *56*, 247–254.

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.

Krapp, H.G., and Hengstenberg, R. (1996). Estimation of self-motion by optic flow processing in single visual interneurons. Nature *384*, 463.

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.

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.

Kühn, N.K., and Gollisch, T. (2016). Joint Encoding of Object Motion and Motion Direction in the Salamander Retina. J. Neurosci. *36*, 12203–12216.

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

Longuet-Higgins, H.C., and Prazdny, K. (1980). The interpretation of a moving retinal image. Proceedings of the Royal Society of London. Series B. Biological Sciences *208*, 385–397.

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. Journal of Neuroscience *33*, 13946–13956.

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.

Maaswinkel, H., and Li, L. (2003). Spatio-temporal frequency characteristics of the optomotor response in zebrafish. Vision Research *43*, 21–30.

Marques, J.C., Lackner, S., Félix, R., and Orger, M.B. (2018). Structure of the Zebrafish Locomotor Repertoire Revealed with Unsupervised Behavioral Clustering. Current Biology *28*, 181-195.e5.

Nakayama, K., and Loomis, J.M. (1974). Optical Velocity Patterns, Velocity-Sensitive Neurons, and Space Perception: A Hypothesis. Perception *3*, 63–80.

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.

Portugues, R., Haesemeyer, M., Blum, M.L., and Engert, F. (2015). Whole-field visual motion drives swimming in larval zebrafish via a stochastic process. Journal of Experimental Biology *218*, 1433–1443.

Qian, H., Zhu, Y., Ramsey, D.J., Chappell, R.L., Dowling, J.E., and Ripps, H. (2005). Directional Asymmetries in the Optokinetic Response of Larval Zebrafish (Danio rerio). Zebrafish *2*, 189–196.

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.

Roeser, T., and Baier, H. (2003). Visuomotor behaviors in larval zebrafish after GFP-guided laser ablation of the optic tectum. Journal of Neuroscience *23*, 3726–3734.

Sabbah, S., Gemmer, J.A., Bhatia-Lin, A., Manoff, G., Castro, G., Siegel, J.K., Jeffery, N., and

Berson, D.M. (2017). A retinal code for motion along the gravitational and body axes. Nature *advance online publication*.

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.

Srinivasan, M., Zhang, S., Lehrer, M., and Collett, T. (1996). Honeybee navigation en route to the goal: visual flight control and odometry. Journal of Experimental Biology *199*, 237–244.

Stone, T., Webb, B., Adden, A., Weddig, N.B., Honkanen, A., Templin, R., Wcislo, W., Scimeca, L., Warrant, E., and Heinze, S. (2017). An Anatomically Constrained Model for Path Integration in the Bee Brain. Current Biology *27*, 3069-3085.e11.

Sunkara, A., DeAngelis, G.C., and Angelaki, D.E. (2016). Joint representation of translational and rotational components of optic flow in parietal cortex. PNAS *113*, 5077–5082.

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. Journal of Neurophysiology *62*, 626–641.

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.

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 Biology *17*, 29.

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 Reports *30*, 442-453.e6.

Wang, K., Arrenberg, B., Hinz, J., and Arrenberg, A.B. (2021). Reduction of visual stimulus artifacts using a spherical tank for small, aquatic animals. Scientific Reports *11*, 3204.

Warren, W.H., and Hannon, D.J. (1988). Direction of self-motion is perceived from optical flow. Nature *336*, 162–163.

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.

Wylie, D.R. (2013). Processing of visual signals related to self-motion in the cerebellum of pigeons. Front Behav Neurosci *7*.

Wylie, D.R.W. (2000). Binocular neurons in the nucleus lentiformis mesencephali in

pigeons: responses to translational and rotational optic flowfields. Neuroscience Letters *291*, 9–12.

Wylie, D.R., Kripalani, T., and Frost, B.J. (1993). Responses of pigeon vestibulocerebellar neurons to optokinetic stimulation. I. Functional organization of neurons discriminating between translational and rotational visual flow. J Neurophysiol *70*, 2632–2646.

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

Zhang, Y., and Arrenberg, A.B. (2019). High throughput, rapid receptive field estimation for global motion sensitive neurons using a contiguous motion noise stimulus. Journal of Neuroscience Methods *326*, 108366.