Supplementary Information

Table of Contents

Supplementary Methods

Tomography of vestibular and ocular systems

Four 2-month old C57Bl/6J mice (2 males, 2 females) were purchased post-mortem (Charles River Laboratories, Harlow, UK) and fixed in phosphate buffered formalin. We followed established procedures for high-resolution microCT imaging of soft as well as hard-tissues[43]. Heads were incubated for 10 days in 10% w/v iodine potassium iodide (I2KI) in formalin. Heads were then rinsed, sealed in polyethylene bags and microCT imaged (80-105kV; 85-105uA; 1440 projections) at Manchester University using the X-Tek 320kV open-bay system (Nikon Metrology, Tring, UK). Voxels were reconstructed with vertices between 32 and 37 µm. Three-dimensional Cartesian co-ordinates representing the right and left sides of each mouse head were collected from 3D reconstructions and orthogonal slices in Avizo 8.1 (FEI, Burlington, USA). Each suite contained 35 landmarks (16 bilateral and 3 midsagittal) corresponding to: the centroid of the lens (Lns); center of the junction between the optic nerve and the bulbus oculi (Opi); origin of the medial rectus muscle on the annulus of Zinn (MRo); insertion of the medial rectus onto the bulbus oculi (MRi); origin of the lateral rectus muscle on the annulus of Zinn (LRo); insertion of the lateral rectus onto the bulbus oculi (LRi); center of the anterior canal lumen at the opening of the ampulla (ASCa); center of the anterior canal lumen at the widest point from the body of the vestibule (ASCx); center of the canal lumen equidistant along the common crus, from the vestibule opening to the bifurcation (CC); center of the posterior canal lumen as it passes the lateral canal (PSC+); center of the posterior canal lumen at the widest point from the body of the vestibule (PSCx); center of the lateral canal lumen at the opening of the ampulla (LSCa); center of the lateral canal lumen at the opening of the vestibule (LSCv); center of the lateral canal lumen at the widest point from the body of the vestibule (LSCx); the intersection between coronal and sagittal sutures (bregma, Brg); the intersection between the interparietal lambdoid and sagittal sutures (lambda, Lbd); midpoint between ventral margins of incisal alveoli (Ia). Sets of landmarks representing each individual were co-registered in MorphoJ version 1.05b with a full Procrustes fit aligned to a plane

defined by Ia-Lbd-Brg. The symmetric component of these co-registered data were then used to calculate a series of vectors for the ocular apparatus and vestibular system.

Using microCT data that were acquired without soft-tissue staining and for isolated otic capsules, kindly provided by Honda et al. [44] for six 42 day old C57BL/6J mice, the otoconial surfaces of the saccular (SM) and utricular (UM) maculae were defined by 4 landmarks marking the maximum length and width. Co-ordinates for these landmarks were then co-registered with the mean of symmetric dataset (see above) using three common landmarks (LSCv, PSC+, and LSCa) and a Procrustes fit while allowing translation, reflection, orthogonal rotation, and scaling (Matlab). A plane for each macula was then defined using a principal components analysis (PAST v2.17c). Whilst these planes are a simplification of the complex geometry of each maculae, they are consistent with omnidirectional sensitivities of these organs and the orthogonal planar relationships reported for Daubenton's Bat (*Myotis daubentonii*) [45]. We determined the collection of 'best axes' that are formed when the head is translated along either the saccula's or utricle's plane. This was achieved by calculating all possible vectors perpendicular to the normal vector of either the saccular or utricular macula.

Because our tomographic data were obtained in fixed tissue, they do not account for the tiltMOR (see Methods); moreover, the fixation and I2KI treatments used for tomography might deform soft tissues[46]. Thus, we could not use tomography to obtain reliable vectors for the optic axis and extra ocular muscles. Instead, the optic axis vector was adopted[40]. See Supplementary Table 2 for vectors of ocular and vestibular elements in a stereotaxic coordinate system.

Supplementary Notes

Supplementary Note 1

We expected ON-DSGCs to exhibit a geometry distinct from that of ON-OFF DSGCs because ON-DSGCs reportedly comprise three subtypes[18,19] with cardinal DS preferences at intervals of 120°, not four at 90°. Furthermore, they are the dominant input to the accessory optic system (AOS)[22,42,47,48], which drives image-stabilizing optokinetic reflexes by encoding rotatory optic flow about semicircular canal best axes[24].

Supplementary Note 2

Because amacrine cells were about as abundant as ganglion cells in our imaged sample, the incidence of DSGCs among all imaged cells (6.6% ON-OFF-DSGCs; 1.8% ON-DSGCs) should be roughly doubled to obtain an estimate of their incidence among all ganglion cells. By this estimate, ~13% of all RGCs are ON-OFF DSGCs, ~4% are ON-DSGCs, and all identified DSGCs together totaled ~17% of all RGCs. These percentages are consistent with earlier reports [23,49]

Supplementary Note 3

The lower half of the flow-tuning plots represents the visual hemifield, while the upper half represents the 'blind hemifield' of extrapersonal space, not seen by this retina. Hotspots for paired subtypes (N-T; V-D) are offset 180° along the abscissa and reflected about the margin of the visual hemifield (e.g., N- and T-cells; Fig. 2i,j). When a flow field well matched to the sampled DS preferences has a center of *contraction* within the visual field (and thus on the retina), the associated hotspot appears in the plot's lower half (V- and T-cells; Fig. 2h,j; 3a,c). When the favored center of *contraction* lies outside the visual hemifield, the hotspot appears in the plot's upper half (N- and D-cells; Fig. 2i, 3a,c). In that case, the diametrically opposed best center of *expansion* is imaged on the retina at or near the best center of *contraction* for its paired subtype (Fig. 3b).

Supplementary Note 4

Application of the stringent DSI criterion to plots of local direction preference produced 3-lobed distributions, both in the central retina[18,19,50] (Fig. 4d) and more peripherally too (Fig. 4f,g; Extended Data Fig. 7). Though N-type ON-DSGCs were poorly tuned, their response kinetics were characteristic of other ON-DSGC subtypes, and clearly different from N-type ON-OFF-DSGCs (Extended Data Fig. 7).

Supplementary Note 5

Probing translatory-flow-matching subtypes for concordance with *rotatory* optic flows (Extended Data Fig. 8o-s), revealed that a subtype's best rotatory axis was perpendicular to two other axes: its best translatory axis, and the optic axis (Fig. 5e, blue and green, respectively). Therefore, each eye contains 8 rotation-tuned DS ensembles, 4 ON-DSGC and 4 ON-OFF-DSGC subtypes, each preferring one direction of rotation about one of two best axes (Fig. 5f).

All best rotatory axes pass through the margin of the visual hemifield. Unlike best translatory axes, they are not identical for the two eyes, giving four distinct best rotatory axes overall (Fig. 5f).

Supplementary Note 6

The regression analysis used to compare the two models' goodness of fit to the data was performed as follows. The flow tuning map of a given DSGC sample (e.g., all imaged ON-OFF DSGCs) was fit to a linear model comprising four individually weighted DS subtypes. For the *translatory* model, members of each subtype aligned their DS preferences everywhere with the optic flow produced by *translation* along that ensemble's best axis. Coordinates of these four best axes were empirically determined using the concordance index, and correspond to the centers of four hotspots (local maxima) in the *translatory* flow-tuning map for imaged or recorded cells. Similarly, for the *rotatory* model, each of the four subtypes aligned its DS preferences with the optic flow produced by *rotation* about a single best axis, empirically determined by concordance analysis, as before. The multiple correlation coefficient (R^2) used here reflects the degree to which these models accurately predict, cell by cell, the directional preferences we determined empirically.

The translatory model significantly outperformed the best rotatory one for all imaged cells ON-OFF-DSGCs (R^2 = 0.97 vs. 0.92) and ON-DSGCs (R^2 = 0.95 vs. 0.87) and for every other sample tested, including for individual DSGC populations (R^2 = 0.78 vs. 0.22 for patch-recorded Hb9 cells; R^2 = 0.72 vs. 0.43 for patched Trhr-GFP cells; and R^2 = 0.88 vs. 0.65 for CART-negative imaged cells). The translatory model particularly outperformed the rotatory one in the temporal periphery, where the directions these models predict are most divergent (Extended Data Fig. 8t-w).

We used bootstrapping to estimate the two-tailed 95% confidence intervals of R^2 values. For each DSGC sample, we resampled the actual DSGCs with replacement, generated a *translatory* flow-tuning map, fit that map with the four individually weighted single-subtype maps comprising the *translatory* model, and determined the multiple correlation coefficient (R^2) value that represents how well the model fits the data. We repeated this procedure 1000 times and calculated the confidence intervals of R^2. Confidence intervals were then determined in the same way for the *rotatory* model.

For all samples tested, the R^2 confidence intervals (CI) for the translatory and rotatory models did not overlap, and the R^2 value obtained for the translatory model was significantly larger ($p < 0.05$) than that obtained for the rotatory model [imaged ON-OFF-DSGCs (CI_{trans}: 0.969-0.971, CI_{rot}: 0.917-0.922), imaged ON-DSGCs (CI_{trans}: 0.947-0.953, CI_{rot}: 0.865-0.875), Hb9 V-cells (CI_{trans}: 0.760-0.799, CI_{rot}: 0.206-0.235), Trhr N-cells (CI_{trans}: 0.698-0.743, CI_{rot}: 0.412-0.447), CART-negative T-cells (CI_{trans}: 0.866-0.894, CI_{rot}: 0.631-0.669)].

Additionally, to add support to this conclusion, we selected the 10 (out of 26) retinas that included the largest number of imaged DSGCs. We then calculated the flow tuning map of each retina separately, for either translatory or rotatory geometry. In nine of them (probability = 0.00977), the translation model fit the DS preferences of DSGCs better (higher R^2) than the rotation model. Moreover, the R^2 for translatory and rotatory geometry differed significantly

(randomization test, p = 0.0272, n_{trans} = 10, n_{rot} = 10). We therefore conclude, based on several lines of evidence, that translatory geometry is superior to rotatory geometry in explaining the DS preferences of DSGCs.

Supplementary Note 7

MRo, MRi, LRo, LRi: origins and insertion points of medial and lateral recti. ASC, PSC, LSC: anterior, posterior and lateral semicircular canals; **X**SCx, center of lumen at apex (point furthest from vestibule); **X**SCa, center of canal ampulla; PSC+, center of posterior canal as it passes LSC; CC, center, bony canal of common crus; SM,UM, saccular and utricular maculae.

Supplementary Note 8

Tomographic reconstruction of the vestibular system revealed that the two cardinal translatory visual axes lay nearly within the two otolithic planes (saccular and utricular; Fig. 5j, left), though many other translatory axes are also represented within these otolithic planes. For rotation, by contrast, there are four best axes among DSGCs, but six among the hair cells of the canals.

Supplementary Note 9

Direction-selectivity index (DSI) was higher for grating stimuli than for bars, especially for ON-DSGCs (ON-DSGCs: $DSI_{gratings}$=0.51 [0.44, 0.56] median [1st quartile, 3rd quartile], DSI_{bars}= 0.12 [0.11, 0.16]; t=56.21, n=197, p<0.001; ON-OFF-DSGCs: $DSI_{gratings}$=0.45 [0.40, 0.50], DSI_{bars}=0.13 [0.11, 0.16], t=98.08, n=662, p<0.001).

Supplementary Note 10

The great majority of CART-positive DSGCs were classified as ON-OFF-DSGCs (353/394; 90%). It is uncertain whether the rest were ON-DSGCs misclassified as ON-OFF-DSGCs based on their bar responses, or were, instead, ON-DSGCs that unexpectedly expressed CART. Though the latter possibility has not been previously suggested [22,23], it is supported by evidence from the Allen Mouse Brain Connectivity Atlas (http://connectivity.brain-map.org) that CART-expressing ganglion cells innervate the MTN[41], which is thought to receive only ON-DSGC input[22]. Similarly, though most RGCs retrolabeled from the AOS were classified as ON-DSGCs, a minority were classified as ON-OFF-DSGCs, including 33% of cells retrolabeled from the superior fasciculus of the optic tract (SF-AOT; 10/30), 24% of those retrolabeled from the inferior fasciculus (IF-AOT; 25/33), and 23% of those retrolabeled from the dorsal division of the medial terminal nucleus (MTNd; 37/48, 77%).

Supplementary Note 11

Throughout this report, we construct models of this form in two steps. We first find the translatory axes that produce optic flow best aligned with the DS cell sample in question, and use these axes to model single-subtype ensembles with each cell in the ensemble matching its DS preferences everywhere to optic flow produced by translation in one of these directions. We then determine what relative weighting of these basis functions, after linear summation, provides the best fit to the real data (a measure of the relative abundance of each subtype in the sample). To test the predictive power of our model, we randomly divided the ON-OFF-DSGC dataset into two halves, a training set (flow tuning plot in Extended Data Fig. 5**k**) and a test set (5**l**). We

determined four best translatory axes for the training set, and then asked how well a weighted linear model of the usual form based on these axes could predict the directional preferences in the test set. **m**, The flow-tuning plot for the cells modeled in this way closely matched that plot for the actual data for cells in the test set (**k**; R^2=0.95). Thus, models of this form are useful as more than compact descriptions of directional data of this form; they have predictive power for other samples of the same type studied under our experimental conditions. **n**, As a further test of predictive power, we asked how well a model using translatory axes derived from the ON-OFF-DSGCs dataset could predict the flow patterns in a completely different population of cells, the imaged ON-DSGCs cells, after optimizing the weighting coefficients for this sample. In this case, too, the model captured the structure of the ON-DSGC flow-tuning plot (R^2=0.83) and recapitulated the apparent weighting of subtypes extracted from direct modeling of ON-DSGC (Fig. 4j,m). Thus, models of this form appear broadly applicable to, and predictive of, the behavior of diverse DSGC populations.

Supplementary Note 12

For ON DSGCs, DSI differed significantly among subtypes (F=22.28, d.f.=3, n=492, p<0.001, bootstrap samples=5000). Post-hoc pair-wise comparisons revealed that DSI of N-type ON-DSGCs was significantly lower than those of the other three ON-DSGC subtypes (t=6.59, 8.67, and 6.78 for the D-, T- and V-cells, p<0.001, bootstrap samples=5000).

For ON-OFF DSGCs, by contrast, applying a more stringent DSI criterion did not relative abundance of ON-OFF-DSGC subtypes (Extended Data Fig. 7e-h). Increasing the DSI threshold from 0.3 to 0.5 decreased the number of DS cells in the sample roughly uniformly among subtypes (77-84%), suggesting little difference among them in selectivity. However, as for ON-DSGCs, ON-OFF DSGC subtypes differed significantly in DSI (F=8.72, d.f.=3, n=1952, p=0.019, bootstrap samples=5000). Post-hoc pair-wise comparisons revealed that DSI of N- and V-type ON-OFF DSGCs was significantly higher than those of T- and V-cells (N-cells: t=3.23 and 3.78 for the D and T subtypes, p<0.005; V-cells: t=3.16 and 3.54 for the D and T subtypes, p<0.005; bootstrap samples=5000). The distribution of DSI values among each subtype deviated significantly from normal distribution (Kolmogorov-Smirnov test) and the distribution of DSIs differed significantly among subtypes (Levene's test). Therefore, to test the effect of subtype on DSI, we used bootstrapping ANOVA routines that do not assume normality or homogeneity of variance. The significance level for pair-wise post-hoc tests was adjusted following the Bonferroni correction for multiple comparisons.

Latency to ON peak (Extended Data Fig. 7k,l), measured from estimated time of arrival bar edge at receptive field, differed significantly between each ON-DSGC subtype and its matching ON-OFF-DSGC subtype (N-cells: t=15.14, d.f.=902, p<0.0001; D-cells: t=21.34, d.f.=488, p<0.0001; T-cells: t=17.91, d.f.=428, p<0.0001; V-cells: t=18.25, d.f.=619, p<0.0001; bootstrap samples=5000). Slope of decay from the ON peak (Extended Data Fig. 7m,n) differed significantly between each ON-DSGC subtype and its matching ON-OFF-DSGC subtype (N-cells: t=23.73, d.f.=902, p<0.0001; D-cells: t=17.38, d.f.=488, p<0.0001; T-cells: t=19.74, d.f.=428, p<0.0001; V-cells: t=13.18, d.f.=619, p<0.0001; bootstrap samples=5000).

Supplementary Note 13

For rotatory optic flow tuning plots (e.g. Extended Data Fig. 8a,b), by convention, the axis corresponding to leftward head rotation a vertical axis (yaw left) is plotted in the inferior visual field, and that for leftward rotation around the body axis (roll left) lies in the anterior visual field. Each of the four hot bands in the plot corresponds to one DSGC subtype. Their ~90° x-axis displacement from one another around the retinal margin reflects the orthogonality of the best rotatory axes.

For modeled rotatory-optic-flow matching DSGCs, cardinal rotatory axes were derived from the centroids of the hotspots in (Extended Data Fig. 8**a**). Rotatory-optic-flow matching subtypes are named for their preferred direction of motion in the central retina. Though modeled rotatory cells thus prefer roughly the same directions as their corresponding translatory subtype in the central retina, marked discrepancies arise near their respective singularities (centers of expansion or contraction).

Supplementary Note 14

This map in Extended Data Fig. 9a is analogous to flattened translatory-flow-tuning maps based on the concordance index (e.g., Fig. 2g-i). However, whereas those plot the percentage of cells aligned within 10º of a specific translatory flow field, here we plot the pooled spike output of the entire ensemble of modeled cells. Each cell's spike output was modeled by using the function in (**a**) to convert the difference angle between predicted and observed preferred directions into a normalized response amplitude. Other models in this report, used to explain the patterns observed among our recorded or imaged cells, matched the locations of modeled cells to the recorded or imaged ones. Because these surely introduce distortions in the global pattern, here each modeled subtype was distributed uniformly across the retina, rather than matched to real cell locations.

Flow tuning plots of the alternate forms in Extended Data Fig. 9b and c have complementary advantages. The thresholding imposed by the concordance index (**c**) sharpens the apparent tuning and provides a better estimate of the best axis of translation, whereas the ensemble-output index (**b**) provides a more accurate reflection of actual gradations of globally summed subtype activity evoked by any possible translatory optic flow. Best axes for each subtype in Extended Data Fig. 9d were derived from data from ON-OFF-DSGCs because it is the largest single sample of DSGCs available; however, axes derived from ON-DSGCs would have differed only modestly.

Hotspots in the maps for translation (Extended Data Fig. 9e, first and third columns) indicate the best egocentric direction of heading for activating the ensemble of DS cells of that subtype Alternatively, one can view the hotspots as indicating the centers of expansion of the optimal translatory flow field, which lies almost exactly straight ahead for N- (advance) cells, behind for T- (retreat) cells, above for D- (rise) cells, and below for V- (fall) cells. Note that due to the distortions inherent in equirectangular cartographic projections (see Fig. 2j,k), points directly above or below the animal are smeared out in the form of lines marking the upper and lower borders of these plots. The response of the same modeled cells to global image *rotation* are shown in the second and fourth columns. Hot spots correspond to the location in egocentric space of the center of counterclockwise circulation of rotatory optic flow as seen from the mouse's perspective (roughly overhead, but slightly into the contralateral hemifield, for N-cells; roughly ahead but slightly into the contralateral hemifield for V-cells, and so forth). These best

translatory and rotatory axes are shown in schematic three-dimensional form in relation to the animal's extrapersonal space in Figs. 5a-d and 5f.

For the self-motion decoder (Extended Data Fig. 9f,g), patterns of relative activation of subtypes by translatory or rotatory flow were drawn from data in Extended Data Fig. 9e. In each plot in the left column of Extended Data Fig. 9f, only one axis of translation produces a pattern of relative activation of the eight DSGC subtypes that matches the pattern evoked by translation along the unknown input axis, and that matching axis corresponds precisely to the input axis. Similarity in the pattern of channel activation falls off (gradually warmer colors) with increasing deviation of a test axis from the reference axis. For the right column of plots in Extended Data Fig. 9f, the coldest spot marks the orientation of the axis of *rotation* producing an 8-channel output pattern most similar to that produced by translation along the input (unknown) axis. This best rotatory axis lies 90° from input axis and thus from the cold spot for the top left plot, as expected (see Fig. 5e). However this cold spot is not as deep a shade of blue as that for the best translatory axis, that is, the minimum Euclidean distance (listed at the top right of each plot) is not as low. This reflects the fact that any translatory-flow-matching cell ensemble will be better fit by the best-fitting translatory-flow template than by any rotatory one. A decoder could use this difference in minimum Euclidean distance to correctly infer that the reference (unknown) self-motion is a translation through space, rather than a rotation.

We implemented in software a simple decoder that was invariably successful in recovering the identity and orientation of any arbitrarily selected translatory or rotatory receptive field simply by finding the most similar 8-channel pattern among a lookup table of all such patterns for all possible rotations and translations. The significance of the foregoing theoretical analysis is that the relative activation of the multiple translatory DS channels of murine DSGCs, pooled across each retina, provides sufficient information to the brain to permit it to describe the mouse's own motion in space. We do not argue that this capacity is unique to the translatory-flow-matching model developed here. Many alternative models comprised of multiple motion sensor subtypes with distinct directional preferences (such as a rotatory flow-matching model) could permit flawless decoding as well. For this reason, the foregoing analysis should not be viewed as lending weight to our argument that murine DSGCs match their DS preferences to cardinal translatory (and not to rotatory) optic flow fields.

Supplementary Note 15

The model of postsynaptic cells tuned for translatory or rotatory optic flow exploits three organizational principles. 1) Best translatory axes are shared by subtype ensembles of the same name in the two eyes (e.g. left-eye and right-eye N-cells or advance cells; Fig. 5b). Cells of the same type in the two eyes can thus be combined into a single panoramic translatory-tuned ensemble by postsynaptic convergence (T1-T4; cf. Fig. 5c,d). 2) Each subtype ensemble is paired with a second ensemble that prefers the opposite direction of motion along the same translatory axis. The difference between these signals is presumably more informative about the animal's direction of translation than either subtype signal in isolation. 3) Whereas homonymous channels in the two eyes prefer the *same* direction along a single axis of *translation*, they prefer *opposite* directions of *rotation* around an orthogonal axis (R1-R6). Thus, for example, a channel built by summing the N-cell signals from the left and right eyes will prefer forward translation

over any rotation (T1) whereas subtracting the N-cell signal from the left eye from that from the right eye yields cells preferring leftward rotation over any translation (R1).

Supplementary Discussion

Relationship of DS subtypes in this study to previously reported subtypes.

The diversity of DS types in mouse retina has grown significantly in recent years. The canonical ON-DSGC and ON-OFF-DSGC varieties[14,15,21,22,51], closely resemble their presumed counterparts in classic rabbit studies[4,18,52-55]. But new non-canonical ON-, and ON-OFF-, and OFF-DSGC types have emerged [15,35,49,56,57]. Which of these subtypes were included in our sample of imaged DS?

Canonical ON-OFF-DSGCs - Overall, our sample of imaged DSGCs certainly includes all four of the canonical ON-OFF-DSGC types in mice; indeed, these probably account for most of our imaged ON-OFF-DSGCs. Among a small sample of such cells targeted for intracellular dye filling, all had dendritic-field sizes and branching patterns typical of the canonical types, and similar stratification and co-fasciculation with starburst amacrine-cell dendrites, which are the key inhibitory element in the classical DS circuit. In the central retina, our imaged cells appear to exhibit the same cardinal preferred directions as the canonical ON-OFF-DSGC types. And they exhibit the same pattern of expression of various genetic or immunohistochemical markers. For example, like canonical murine ON-OFF-DSGCs, three of the ON-OFF-DSGCs in our sample (N-, V-, and D-cells) are CART-immunopositive, while the fourth (T-cells) was CART-immunonegative. When two of the canonical ON-OFF-DSGC types (V- and N-cells) were targeted for patch recording using transgenic reporter mice (Hb9-GFP; Trhr-GFP; Fig. 2a,b), their DS preferences exhibited the same topographic pattern of DS preference as the corresponding subtype of imaged ON-OFF-DSGCs.

We were intrigued to discover that the increased abundance of N-type ON-OFF-DSGCs relative to the other subtypes in our data closely matched those originally reported by Oyster and Barlow in rabbits[18], with N-type ON-OFF-DSGCs twice as common as other subtypes (Supplementary Note 9; current vs. original study; N-cells: 41 vs. 44, D-cells: 16 vs. 11, T-cells: 17 vs. 22, and V-cells: 25 vs. 23). This may be linked to the statistics of the animal's self-motion. Forward ambulation, a common occurrence in mice, as in most bilaterally symmetric animals, will preferentially activate N-cells. Caution is in order here, however, both because cross-species comparisons are tricky, and because an earlier study[58] found N- and T-cells to be about equally abundant in adult mice.

Both Trhr-GFP and Drd4-GFP reporter mice exhibit selective GFP labeling of N-type (posterior-motion-preferring) ON-OFF-DSGCs[15]. However, the tagged cells in the two strains differ in several properties, suggesting that N-cells might be formally divisible into two subtypes. If so, both apparently adhere to the translatory spherical geometry we have mapped. We show this for Trhr-GFP DSGCs in Fig. 2b. Pilot patch-recorded data from Drd4-GFP cells are also consistent with this geometry (data not shown).

Canonical ON-DSGCs - The other canonical DS class — the ON-DSGCs — are well established as the dominant source of retinal input to the AOS. Cells we retrolabeled from the AOS were independently and blindly classified as ON-DSGCs based on our calcium-imaging data (Extended Data Fig. 2g-k, Extended Data Fig. 3i-m). Significantly, even the N-type of ON-DSGC, which has apparently been missed in earlier studies, is represented among cells labeled by retrograde transport from the AOS. Cells we classified as ON-DSGCs based on calcium imaging were overwhelmingly CART-immunonegative, as previously reported for classic ON-DSGCs[22].

Novel DSGC types - There are two new additions to the ON-DS class in mice: the transient ON-DSGCs recently shown to innervate the colliculus[57] and the F-mini^{ON} cells[56] . Both of these have morphologies clearly distinct from those of the classic ON-DSGC type[22]; a similar story has emerged in rabbit retina[19,55]. We did not encounter cells of this form among those we filled with dye after identifying them as ON-DSGCs by calcium-imaging, but our sample is small. Thus, while such novel types could be among our sample of imaged ON-DSGCs, we have no direct evidence that they are. In both mice[57] and rabbits[19,55] the novel transient ON-DS cells comprise the same set of directional subtypes as the canonical AOS-projecting "sustained ON-DS" cells. If so, they seem likely to follow the same global geometry we have documented for the canonical DSGCs. This has implications for the underlying retinal synaptic circuitry. Neither the F-mini^{ON} cells in mice nor the transient ON-DS cells of rabbit costratify with starburst amacrine cells and my not depend on them for the computation of direction. This raises the possibility that the translatory-optic-flow geometry of DS preference outlined is not limited to the canonical DS types and the starburst networks that shape them, but may extend to non-canonical DS types and the novel amacrine-cell circuits that presumably mediate their selectivity.

We did not encounter OFF-DSGCs in our imaging studies, though there are apparently at least two types of such cells, the JAM-B cells[34,59] (Kim) and F-mini^{OFF} cells.[49] (see also[36]). This may be because our moving stimuli were designed to optimally activate canonical DSGCs; their dimensions and velocities differed from those in studies detecting OFF-DSGCs in mouse retina[35,49,56]. At least one of these novel types, the JAM-B cells, have a low direction-selectivity index (<0.3), which drops even lower (0.1) in a moderately light background[59], like that used in our imaging experiments. Thus, it remains to be determined whether these OFF-DSGC subtypes adhere to the spherical geometry we describe for to the canonical types. Again, the circuitry implications are intriguing; these subtypes apparently derive their DS preference in part from dendritic asymmetries, and not from synaptic interactions with starburst amacrine cells.

Possible misidentification of canonical ON- and ON-OFF DSGCs

The unexpected similarity between ON- and ON-OFF-DSGCs (cf. Figs. 1f and 3c) raised the concern that we had misclassified many ON-OFF-DSGCs as ON-DSGCs; a related concern was that nearly all cells we classified as ON-DSGCs unexpectedly exhibited OFF voltage or calcium responses. However, the polar plot remained cruciform in a much smaller sample of ON-DSGCs cells meeting a particularly stringent classification criterion (>95% posterior probability) (Extended Data Fig. 6). Additionally, we studied how the two response features used for unsupervised clustering (time to ON peak and the slope following the ON peak) vary among the

four translation channels and among ON- and ON-OFF-DSGCs. This analysis demonstrated the occurrence of 4 distinct ON-DSGC subtypes and 4 distinct ON-OFF-DSGCs subtypes (Extended Data Figure 7i-n). The unexpected OFF voltage and calcium responses proved to be a normal feature of murine ON-DSGCs. They were readily detected in cell samples largely or completely restricted to ON-DSGCs by considering only: 1) patched and dye-filled cells with canonical ON-DSGC morphology (Extended Data Fig. 1e-i, Extended Data Fig. 3i); or 2) only cells retrolabeled from the MTN, which derives its input overwhelmingly if not exclusively from ON-DSGCs [22,42] (Extended Data Fig. 3h). In retrospect, OFF excitation in ON-DSGCs is not surprising because virtually all of them have a minor dendritic arbor in the OFF cholinergic band [22]; this OFF excitation persisted when blocking the ON synaptic transmission (Extended Data Fig. 1i). Together, these evidence suggest that the probability of misclassification of ON- and ON-OFF-DSGCs is minimal.

Supplementary Equations

1. **Transformation from flat mounted retina to spherical retina**

The retina was modeled as a spherical object S of radius R with maximal longitudinal distance M. Specifically, we assumed a domain of parametrization $\Omega = [0, M] \times [0, 360°]$ and parametrized S by a map $\mathbf{x} : \Omega \mapsto S \subset \mathbb{R}^3$ defined by

$$\begin{aligned}\mathbf{x}(s,\theta) &= (x(s,\theta), y(s,\theta), z(s,\theta)), \\ &= \left(R\sin\left(\frac{s}{R}\right)\cos(\theta), R\sin\left(\frac{s}{R}\right)\sin(\theta), R - R\cos\left(\frac{s}{R}\right)\right),\end{aligned} \tag{S1}$$

where s/R is the co-latitude coordinate as measured from the optic disc and θ is the longitudinal coordinate measured from the line segment connecting the optic disc to the insertion point of the lateral rectus muscle. Four relieving cuts in the retina were assumed to lie along meridians (Extended Data Figure 4f, red lines) dividing Ω into sectors $\Omega_i = \{(s,\theta) \in \Omega : \theta_i < \theta < \theta_{i+1}\}$ where θ_i denotes the longitudinal coordinate of the i-th cut. The arc length of the circular arc connecting the optic disc to the i-th cut was denoted by m_i (Extended Data 4e,f). The flat-mounted retina was modeled by a domain $\mathcal{D} \subset \mathbb{R}^2$ with Cartesian coordinates $(u, v) \subset \mathcal{D}$. The flattening of the retina was modeled by $F : \Omega \mapsto \mathcal{D} \subset \mathbb{R}^2$ which in $(u, v) \in \mathcal{D}$ coordinates can be described by $F(s,\theta) = (u(s,\theta), v(s,\theta))$ (Extended Data 4c). To be precise, the mapping from the spherical retina to the flat mounted retina is $F \circ \mathbf{x}^{-1}$. However, all relevant quantities for our analysis could be expressed in terms of the intrinsic coordinates (s,θ). In our analysis, we expressed F in a polar coordinate representation of the form $u(s,\theta) = \rho(s,\theta)\cos(f(s,\theta))$ and $v(s,\theta) = \rho(s,\theta)\sin(f(s,\theta))$ where $\rho : \Omega \mapsto \mathbb{R}^+$ and $f : \Omega \mapsto [0, 360°]$ are the radial and angular polar coordinates as measured from $F(0,0)$, i.e. the image of the optic disc under F, and the line segment connecting the optic disc to the insertion point of the lateral rectus muscle. To account for the cuts we defined ρ, f in a piecewise manner: $\rho(s,\theta) = \rho_i(s,\theta)$, $f(s,\theta) = f_i(s,\theta)$ if $(s,\theta) \in \Omega_i$.

To determine the mapping F from the retina into the flat-mounted retina (Extended Data Figure 4c,e) we assumed an elastic model. Specifically, we assumed that the retina is an isotropic elastic material with a linear stress-strain constitutive relationship with the elastic energy of a flattening map given by

$$E[F] = \sum_{i=1}^{4} \int_{\Omega_i} \left[\nu\left(\gamma_{11} + \gamma_{22}\right)^2 + (1-\nu)\left(\gamma_{11}^2 + 2\gamma_{12}^2 + \gamma_{22}^2\right)\right] \sin\left(\frac{s}{R}\right) ds d\theta, \tag{S2}$$

where ν is the Poisson ratio for the material and the components γ_{ij} of the in-plane strain tensor in each sector are defined by:

$$\gamma_{11} = \left(\frac{\partial \rho_i}{\partial s}\right)^2 + \rho_i^2 \left(\frac{\partial f_i}{\partial s}\right)^2 - 1, \tag{S3}$$

$$\gamma_{12} = R^{-1} \sin\left(\frac{s}{R}\right)^{-1} \left(\frac{\partial \rho_i}{\partial s}\frac{\partial \rho_i}{\partial \theta} + \rho^2 \frac{\partial f_i}{\partial s}\frac{\partial f_i}{\partial \theta}\right), \tag{S4}$$

$$\gamma_{22} = R^{-2} \sin\left(\frac{s}{R}\right)^{-2} \left(\left(\frac{\partial \rho_i}{\partial \theta}\right)^2 + \rho_i^2 \left(\frac{\partial f_i}{\partial \theta}\right)^2 - R^2 \sin^2\left(\frac{s}{R}\right)\right). \tag{S5}$$

We further assumed that the retina is a perfectly incompressible material and hence we took $\nu = .5$. To ensure continuity, we assumed upon flattening the following boundary conditions:

$$\begin{cases} \rho_i(s, \theta_{i+1}) = \rho_{i+1}(s, \theta_{i+1}), & 0 \le s \le m_{i+1} \\ f_i(s, \theta_{i+1}) = f_{i+1}(s, \theta_{i+1}) = \theta_{i+1}, & 0 \le s \le m_{i+1} \end{cases}. \tag{S6}$$

The realized flattening map F was taken to be the minimizer of Eq. (S2) subject to the boundary conditions (S6). That is, given the exact values of m_i, θ_i, M and R we assumed the resulting flat mounted retina is described by the map F which minimizes Eq. (S2) and in particular the spherical retina can be reconstructed from the flat mounted retina by the mapping $\mathbf{x} \circ F^{-1}$.

The exact values of m_i, θ_i, M and R could not be accurately measured on the eyeball before dissection. Instead, these quantities were approximated using the following measurements on the flat mounted retina: (1) angular polar coordinates $\overline{\theta}_i$ of the terminal point of the cuts relative to the insertion point of the line segment connecting the optic disc to the insertion point of the lateral rectus muscle, (2) length $\overline{m}_i$ of the line segment connecting the optic disc and the terminal point of the cuts (Extended Data Figure 4a green lines), and (3) average length $\overline{M}$ between the optic disc and the retina's rim, accounting for folding, as measured along lines bisecting the terminal points of cuts (Extended Data Figure 4a blue lines). Specifically, the values of θ_i and m_i were determined from the approximation $\theta_i \approx \overline{\theta}_i$ and $m_i \approx \overline{m}_i$ which is valid provided the cuts are sufficiently deep so that $m_i/R \ll 1$. Indeed, for $s/R < m_i/R \ll 1$ it follows that the nonlinearities in Eqs. (S2-S6) resulting from the spherical geometry are negligible since $\sin(s/R) = s/R + \mathcal{O}(m_i^3/R^3)$. Consequently, for $s/R < m_i/R$ the minimizer of Eqs. (S2-S6) can be well approximated by $\rho_i = s$ and $f_i = \theta$ in each sector. Second, we made the approximation $\overline{M} \approx M$, i.e. the bisecting lines remain straight upon flattening and retain their arc-length. This assumption is valid provided the cuts have approximately the same length,

i.e. $\overline{m}_i \approx \overline{m}_{i+1}$, which implies that the flattened sectors are symmetric about the lines bisecting the terminal points of cuts. Consequently, to retain symmetry, it follows that upon flattening the circular arcs bisecting the cuts on the spherical retina will be mapped to straight lines on the flattened retina. Moreover, the bisecting lines approximately retain their length upon flattening since $\rho_i \approx s$ near the optic disc while the folded parts of the retina are not constrained to remain flat and can exactly match the arc length. Finally, the value of R was determined by the relationship $M/(\pi R) \times 180^\circ = \arccos(1 - 2r/D)$ where D was the measured diameter of the eyeball and r was the measured distance from the optic disc to the projection of the rim of the retina onto the optical axis.

Using the calculated values of m_i, θ_i, M and R we then numerically minimized Eqs. (S2-S6) to approximate the flat mounted retina. To be specific, in each sector Ω_p we used a 40×40 uniform discretization $[s^{i,j}, \theta^{i,j}]$ with spacing $\delta s = M/40$, $\delta\theta = (\theta_{p+1} - \theta_p)/40$. The integral in Eq. (S2) was approximated over a coordinate rectangle $B_{i,j}$ using a composite quadrature rule:

$$E^{i,j}[F] = \frac{\delta r \delta\theta}{4} \sum_{l=i}^{i+1} \sum_{k=j}^{j+1} \left[\nu \left(\gamma_{11}^{l,k} + \gamma_{22}^{l,k} \right)^2 + (1-\nu) \left(\left(\gamma_{11}^{l,k} \right)^2 + 2 \left(\gamma_{12}^{l,k} \right)^2 + \left(\gamma_{22}^{l,k} \right)^2 \right) \right] \sin\left(\frac{s^{l,k}}{R} \right). \tag{S7}$$

The partial derivatives appearing in the strain terms (S3-S5) were then approximated by fourth order finite differences with appropriate forward and backward schemes near the boundary of each sector. Summing over i and j the elastic energy in the sector was approximated by

$$E[F] = \sum_{i=1}^{40} \sum_{j=1}^{40} E^{i,j}[F] + \mathcal{O}(\delta s^2 + \delta\theta^2). \tag{S8}$$

In sector Ω_p, the continuity conditions (S6) between sectors was enforced by assuming

$$\begin{cases} \rho_p^{i,1} = s^{i,1} \text{ and } f_p^{i,1} = \theta_p, & s^{i,1} \le m_p \\ \rho_p^{i,40} = s^{i,40} \text{ and } f_p^{i,40} = \theta_{p+1}, & s^{i,40} \le m_{p+1} \end{cases} \tag{S9}$$

which is valid if $m_p/R \ll 1$. Finally, the discrete sum of quadratic terms in Eq. (S2) was directly minimized using Matlab's minimization routine lsqnonlin subject to the boundary conditions (S9). The numerically computed minimum F was then used to generate a mesh of points $(u^{i,j}, v^{i,j}) = F(s^{i,j}, \theta^{i,j})$ which approximate the location of points on the flat mounted retina (Extended Data 4c). The empirical coordinates $(\overline{u}, \overline{v}) \in \overline{\mathcal{D}}$ of each cell location on the flat mounted retina were then mapped to the mesh $(u^{i,j}, v^{i,j})$ by proximity to a grid point. This

interpolation generates a mapping back to the spherical coordinates denoted by $\overline{F}^{-1}$ and the mapping to the numerical approximation to the retina is through composition $F \circ \overline{F}^{-1}$. Furthermore, the location of the cells on the spherical retina could also be determined by the mapping $\mathbf{x} \circ \overline{F}^{-1}$. This numerical scheme for reconstructing the flat mounted retina qualitatively recapitulates the data in the original flat-mounted experimental retina (Extended Data Figure 4a,c).

To display data pooled across retinas we mapped all cells to a "standard flat mounted retina". This was constructed by numerically minimizing the elastic model (S2-S6) assuming a rim angle of $\phi_0 = 107.50^\circ$ and relieving cuts with $m_i = .4M$ and $\theta_0 = 90^\circ, 180^\circ, 270^\circ, 360^\circ$. The value ϕ_0 was determined from taking the average over all measured values of the rim angle. This generates a flattening map $F' : \Omega \mapsto \mathbb{R}^2$ whose image $F'(\Omega)$ is the standard flat mounted retina with Cartesian coordinates (u', v'). The empirically measured cells on the flat mounted retina were then mapped to the standard flat mounted retina by composition $F' \circ \overline{F}^{-1}$ or equivalently from the numerical approximation to the flat mounted by $F' \circ F^{-1}$ (Extended Data 4c,d).

The direction preference of DSGC cells was mapped to vectors on the different representations of the retina by applying the Jacobian matrix for the various mappings. Specifically, direction preference of a DSGC cell was modeled by a vector $\mathbf{W} = W_1\mathbf{e}_u + W_2\mathbf{e}_v$ with $\mathbf{e}_u$, $\mathbf{e}_v$ the standard (u, v) directions. The Jacobian matrices J, J' for the mappings F, F', respectively, were computed using a fourth order finite difference approximation to the following matrices:

$$J = \begin{pmatrix} \frac{\partial u}{\partial s} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial s} & \frac{\partial v}{\partial \theta} \end{pmatrix} \text{ and } J' = \begin{pmatrix} \frac{\partial u'}{\partial s} & \frac{\partial u'}{\partial \theta} \\ \frac{\partial v'}{\partial s} & \frac{\partial v'}{\partial \theta} \end{pmatrix}. \tag{S10}$$

The vector $\mathbf{W}$ was then mapped to a vector $\mathbf{T} = T_1\mathbf{e}_s + T_2\mathbf{e}_\theta$, $\mathbf{W}' = W_1\mathbf{e}_{u'} + W_2\mathbf{e}_{v'}$ in (s, θ) and (u', v') coordinates by matrix multiplication: $\mathbf{T} = J^{-1}\mathbf{W}$ and $\mathbf{W} = J'J\mathbf{W} = J'\mathbf{T}$. Pooling the data across all retinas we then generated a vector field, also denoted by $\mathbf{W}'$, on the standard retina. The direction preference was also mapped to a vector $\hat{\mathbf{T}} = \hat{T}_1\mathbf{e}_x + \hat{T}_2\mathbf{e}_y + \hat{T}_3\mathbf{e}_z$ on the spherical retina by the mapping

$$\begin{pmatrix} \hat{T}_1 \\ \hat{T}_2 \\ \hat{T}_3 \end{pmatrix} = \begin{pmatrix} \cos(s/R)\cos(\theta) & -R\sin(s/R)\sin(\theta) \\ \cos(s/R)\sin(\theta) & R\sin(s/R)\cos(\theta) \\ \sin(s/R) & 0 \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}. \tag{S11}$$

2. **Mapping of optic flow to the retina**

Translational and rotational optic flow in extrapersonal space was modeled by vector fields on a sphere. Specifically, we denoted co-latitude coordinates $\hat{\phi} \in (0, 180^\circ)$ in visual space measured from the south pole represented by the optical axis and longitudinal coordinates measured from the line segment connecting the optic disc and the insertion point of the lateral rectus muscle by $\hat{\theta} \in (0, 360^\circ)$. The optic flow was modeled by $\hat{\mathbf{V}}(\hat{\phi}, \hat{\theta}) = \hat{V}_1(\hat{\phi}, \hat{\theta})\mathbf{e}_{\hat{\phi}} + \hat{V}_2(\hat{\phi}, \hat{\theta})\mathbf{e}_{\hat{\theta}}$ where $\mathbf{e}_{\hat{\phi}}$, $\mathbf{e}_{\hat{\theta}}$ are the natural un-normalized $(\hat{s}, \hat{\theta})$ coordinate vector fields. For a translation optic flow expanding from the the point $(\hat{\phi}_0, \hat{\theta}_0)$ the vector field $\mathbf{V}$ has the components:

$$\hat{V}_1^T(\hat{s}, \hat{\theta}) = \frac{\cos(\hat{\phi}_0)\sin(\hat{\phi}) - \cos(\hat{\phi})\cos(\hat{\theta}_0 - \hat{\theta})\sin(\hat{\phi}_0)}{\sqrt{1 - \Big(\cos(\hat{\phi}_0)\cos(\hat{\phi}) + \sin(\hat{\phi}_0)\cos(\hat{\theta}_0 - \hat{\theta})\sin(\hat{\phi})\Big)^2}}, \tag{S12}$$

$$\hat{V}_2^T(\hat{s}, \hat{\theta}) = -\frac{\sin(\hat{\phi})\sin(\hat{\phi}_0)\sin(\hat{\theta}_0 - \hat{\theta})}{\sqrt{1 - \Big(\cos(\hat{\phi}_0)\cos(\hat{\phi}) + \sin(\hat{\phi}_0)\cos(\hat{\theta}_0 - \hat{\theta})\sin(\hat{\phi})\Big)^2}}. \tag{S13}$$

For a rotational optic flow about an axis passing through the point $(\hat{\phi}_0, \hat{\theta}_0)$, the vector field $\mathbf{V}$ has the components:

$$\hat{V}_1^R = \frac{\sin(\hat{\phi}_0)\sin(\hat{\theta}_0 - \hat{\theta})}{\sin^2(\hat{\phi})\sin^2(\hat{\theta}_0 - \hat{\theta}) + \Big(\sin(\hat{\phi}_0)\cos(\hat{\phi})\cos(\hat{\theta}_0) - \cos(\hat{\phi})\sin(\sin(\hat{\phi}_0)\Big)^2}, \tag{S14}$$

$$\hat{V}_2^R = \frac{\sin(\hat{\phi})\Big(\sin(\hat{\phi})\cos(\hat{\phi}_0) - \cos(\hat{\phi})\sin(\hat{\phi}_0)\cos(\hat{\theta}_0 - \hat{\theta})\Big)}{\sin^2(\hat{\phi})\sin^2(\hat{\theta}_0 - \hat{\theta}) + \Big(\sin(\hat{\phi}_0)\cos(\hat{\phi})\cos(\hat{\theta}_0) - \cos(\hat{\phi})\sin(\sin(\hat{\phi}_0)\Big)^2}. \tag{S15}$$

To map the optic flow onto the retina we assumed the visual angle for each eye to be 180° and the optics of the eye maps visual space linearly to the entirety of the retina. Specifically, a point $(\hat{\phi}, \hat{\theta})$ in extra personal space was mapped to (s, θ) coordinates on the retina by the linear mapping $s = -2M/180^\circ\hat{\phi} + 2M$, $\theta = \hat{\theta} + 180^\circ$. This mapping induces a vector field $\mathbf{V}$ on the retina via the relationship:

$$\mathbf{V}(s, \theta) = V_1(s, \theta)\mathbf{e}_s + V_2(s, \theta)\mathbf{e}_\theta = -\frac{2M}{180^\circ}\hat{V}_1(\hat{\phi}(s), \hat{\theta}(\theta))\mathbf{e}_s + \hat{V}_2(\hat{\phi}(s), \hat{\theta}(\theta))\mathbf{e}_\theta, \tag{S16}$$

where $\mathbf{e}_s, \mathbf{e}_\theta$ are the natural un-normalized (s, θ) coordinate vector fields. The optic flow was then mapped to a vector field $\mathbf{U}$ on the flat mounted retina and a vector field $\mathbf{U}'$ on the standard flat mounted retina:

$$\begin{aligned}\mathbf{U}(u,v) = &\left(\frac{\partial u}{\partial s}V_1(s(u,v),\theta(u,v)) + \frac{\partial u}{\partial \theta}V_2(s(u,v),\theta(u,v))\right)\mathbf{e}_u \\ &+ \left(\frac{\partial v}{\partial s}V_1(s(u,v),\theta(u,v)) + \frac{\partial v}{\partial \theta}V_2(s(u,v),\theta(u,v))\right)\mathbf{e}_v,\end{aligned} \tag{S17}$$

$$\begin{aligned}\mathbf{U}'(u',v') = &\left(\frac{\partial u'}{\partial s'}V_1(s(u',v'),\theta(u',v')) + \frac{\partial u'}{\partial \theta}V_2(s(u',v'),\theta(u',v'))\right)\mathbf{e}_{u'} \\ &+ \left(\frac{\partial v}{\partial s}V_1(s(u,v),\theta(u,v)) + \frac{\partial v}{\partial \theta}V_2(s(u,v),\theta(u,v))\right)\mathbf{e}_{v'},\end{aligned} \tag{S18}$$

where the coordinate relationships $(u(s,\theta),\, v(s,\theta))$ and $(u'(s,\theta), v'(s,\theta))$ are determined from the numerical model.

The optic flow on the standard flat mounted retina was then quantitatively compared with the direction preference of DSGC's on the standard flat mounted retina. Specifically, denoting the vector field generated from the empirical data by $\mathbf{W}'(u',v')$ the angle ψ between the vector fields at each point was computed by the formula

$$\cos(\psi) = \frac{\langle \mathbf{U}', \mathbf{W}' \rangle_{F'_*\mathbf{g}}}{\|\mathbf{U}'\|_{F'_*\mathbf{g}} \|\mathbf{W}'\|_{F'_*\mathbf{g}}}, \tag{S19}$$

where the inner products and norms are computed under the action of the following matrix:

$$F'_*\mathbf{g} = \begin{pmatrix} \left(\frac{\partial s}{\partial u}\right)^2 + R^2\sin^2(s/R)\left(\frac{\partial \theta}{\partial u}\right)^2 & \frac{\partial s}{\partial u}\frac{\partial s}{\partial v} + R^2\sin^2(s/R)\frac{\partial \theta}{\partial u}\frac{\partial \theta}{\partial v} \\ \frac{\partial s}{\partial u}\frac{\partial s}{\partial v} + R^2\sin^2(s/R)\frac{\partial \theta}{\partial u}\frac{\partial \theta}{\partial v} & \left(\frac{\partial s}{\partial v}\right)^2 + R^2\sin^2(s/R)\left(\frac{\partial \theta}{\partial v}\right)^2 \end{pmatrix}. \tag{S20}$$

Computing the angle via Eq. (S19) is equivalent to mapping the vector fields back to the spherical retina and computing the angle between the vector fields in three dimensions. In particular, Eq. (S19) accounts for the distortion between vectors that occurs when the spherical retina flattened.

Supplementary References

43 Jeffery, N. S., Stephenson, R. S., Gallagher, J. A., Jarvis, J. C. & Cox, P. G. Micro-computed tomography with iodine staining resolves the arrangement of muscle fibres. *J Biomech* **44**, 189-192, doi:10.1016/j.jbiomech.2010.08.027 (2011).

44 Honda, K. *et al.* Ex vivo visualization of the mouse otoconial layer compared with micro-computed tomography. *Otol Neurotol* **36**, 311-317, doi:10.1097/Mao.0000000000000376 (2015).

45 Kirkegaard, M. & Jorgensen, J. M. The inner ear macular sensory epithelia of the Daubenton's bat. *J. Comp. Neurol.* **438**, 433-444, doi:DOI 10.1002/cne.1326 (2001).

46 Vickerton, P., Jarvis, J. & Jeffery, N. Concentration-dependent specimen shrinkage in iodine-enhanced microCT. *J. Anat.* **223**, 185-193, doi:10.1111/joa.12068 (2013).

47 Oyster, C. W., Simpson, J. I., Takahashi, E. S. & Soodak, R. E. Retinal ganglion cells projecting to the rabbit accessory optic system. *J. Comp. Neurol.* **190**, 49-61, doi:10.1002/cne.901900105 (1980).

48 Dann, J. F. & Buhl, E. H. Retinal ganglion-cells projecting to the accessory optic system in the rat. *J. Comp. Neurol.* **262**, 141-158, doi:10.1002/cne.902620111 (1987).

49 Baden, T. *et al.* The functional diversity of retinal ganglion cells in the mouse. *Nature* **529**, 345-350, doi:10.1038/nature16468 (2016).

50 Yonehara, K. *et al.* Congenital nystagmus gene FRMD7 is necessary for establishing a neuronal circuit asymmetry for direction selectivity. *Neuron* **89**, 177-193, doi:10.1016/j.neuron.2015.11.032 (2016).

51 Weng, S., Sun, W. & He, S. Identification of ON-OFF direction-selective ganglion cells in the mouse retina. *J Physiol* **562**, 915-923, doi:10.1113/jphysiol.2004.076695 (2005).

52 He, S. G. & Masland, R. H. ON direction-selective ganglion cells in the rabbit retina: Dendritic morphology and pattern of fasciculation. *Vis. Neurosci.* **15**, 369-375, doi:10.1017/s095252389815215x (1998).

53 Sivyer, B., van Wyk, M., Vaney, D. I. & Taylor, W. R. Synaptic inputs and timing underlying the velocity tuning of direction-selective ganglion cells in rabbit retina. *J. Physiol. (Lond).* **588**, 3243-3253, doi:10.1113/jphysiol.2010.192716 (2010).

54 Taylor, W. R. & Vaney, D. I. Diverse synaptic mechanisms generate direction selectivity in the rabbit retina. *J. Neurosci.* **22**, 7712-7720 (2002).

55 Hoshi, H., Tian, L.-M., Massey, S. C. & Mills, S. L. Two Distinct Types of ON Directionally Selective Ganglion Cells in the Rabbit Retina. *J. Comp. Neurol.* **519**, 2509-2521, doi:10.1002/cne.22678 (2011).

56 Rousso, D. L. *et al.* Two pairs of ON and OFF retinal ganglion cells are defined by intersectional patterns of transcription factor expression. *Cell Rep* **15**, 1930-1944, doi:10.1016/j.celrep.2016.04.069 (2016).

57 Gauvain, G. & Murphy, G. J. Projection-specific characteristics of retinal input to the brain. *J Neurosci* **35**, 6575-6583, doi:10.1523/JNEUROSCI.4298-14.2015 (2015).

58 Elstrott, J. *et al.* Direction selectivity in the retina is established independent of visual experience and cholinergic retinal waves. *Neuron* **58**, 499-506, doi:10.1016/j.neuron.2008.03.013 (2008).

59 Joesch, M. & Meister, M. A neuronal circuit for colour vision based on rod-cone opponency. *Nature* **532**, 236-239, doi:10.1038/nature17158 (2016).