

# Insect-Inspired Estimation of Egomotion

**Matthias O. Franz**

*mof@tuebingen.pg.de*

*Max-Planck-Institut für biologische Kybernetik, Tübingen, Germany*

**Javaan S. Chahl**

*javaan@zappa.anu.edu.au*

*Center of Visual Sciences, Research School of Biological Sciences, Australian National University, Canberra, Australia*

**Holger G. Krapp**

*hgk23@cam.ac.uk*

*Department of Zoology, University of Cambridge, United Kingdom*

**Tangential neurons in the fly brain are sensitive to the typical optic flow patterns generated during egomotion. In this study, we examine whether a simplified linear model based on the organization principles in tangential neurons can be used to estimate egomotion from the optic flow. We present a theory for the construction of an estimator consisting of a linear combination of optic flow vectors that incorporates prior knowledge both about the distance distribution of the environment, and about the noise and egomotion statistics of the sensor. The estimator is tested on a gantry carrying an omnidirectional vision sensor. The experiments show that the proposed approach leads to accurate and robust estimates of rotation rates, whereas translation estimates are of reasonable quality, albeit less reliable.**

## 1 Introduction

A moving visual system generates a characteristic pattern of image motion on its sensors. The resulting optic flow field is an important source of information about the egomotion of the visual system (Gibson, 1950). In the fly brain, part of this information is analyzed by a group of wide-field, motion-sensitive neurons, the tangential neurons in the lobula plate (Hausen, 1993; Egelhaaf, Kern, Krapp, Kretzberg, Kurtz, & Warzecha, 2002). A detailed mapping of their local preferred directions and motion sensitivities (Krapp & Hengstenberg, 1996) reveals a striking similarity to certain egomotion-induced optic flow fields (cf. Fig. 1). This suggests that each tangential neuron extracts a specific egomotion component from the optic flow which may be useful for gaze stabilization and flight steering.

A recent study (Franz & Krapp, 2000) has shown that a simplified computational model of the tangential neurons as a weighted sum of flow measurements was able to explain certain properties of the observed response fields. The weights were chosen according to an optimality principle which minimizes the output variance of the model caused by noise and distance variability between different scenes. In that study, however, we mainly focussed on a comparison between the sensitivity distribution in tangential neurons and the weight distribution of such optic flow processing units. Here we present a classical linear estimation approach that extends the previous model to the complete egomotion problem. We again use linear combinations of local flow mea-

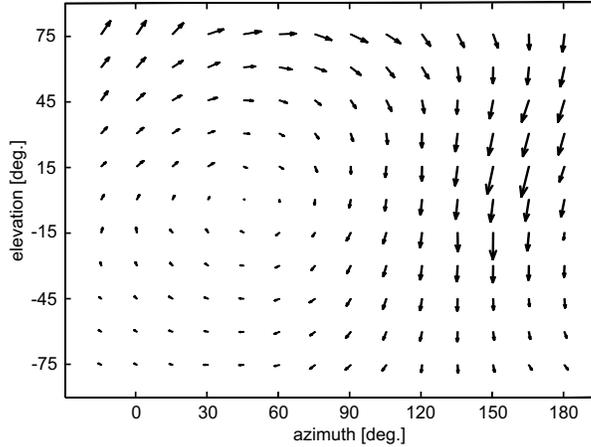


Figure 1: Mercator map of the response field of the neuron VS10. The orientation of each arrow gives the local preferred direction, and its length denotes the relative local motion sensitivity. The results suggest that VS10 responds maximally to rotation around an axis at an azimuth of about  $50^\circ$  and an elevation of about  $0^\circ$  (after Krapp et al., 1998).

surements, but - instead of prescribing a fixed motion axis and minimizing the output variance - we minimize the quadratic error in the estimated egomotion parameters. The derived weight sets for the single model neurons are identical to those obtained from one of the model variants discussed in Franz and Krapp (2000). Of primary interest for this article is, however, that the new approach yields a novel, extremely simple estimator for egomotion that consists of a linear combination of model neurons. Our experiments indicate that this insect-inspired estimator shows - in spite of its simplicity - an astonishing performance that often comes close to the more elaborate approaches of classical computer vision.

This article is structured as follows: in Sect. 2, we describe the derivation of the egomotion estimator from a least squares principle. In Sect. 3, we subject the obtained model to a rigorous real-world test on a gantry carrying an omnidirectional vision sensor. The evidence and the properties of such a neural representation of egomotion are discussed in Sect. 4. A preliminary account of our work has appeared in Franz and Chahl (2003).

## 2 Optimal linear estimators for egomotion

### 2.1 Egomotion sensor and neural model

In order to simplify the mathematical treatment, we assume that the  $N$  motion detectors of our egomotion sensor are arranged on the unit sphere. The viewing direction of the inputs to a particular motion detector with index  $i$  is denoted by the radial unit vector  $\mathbf{d}_i$ . At each viewing direction, we define a local two-dimensional coordinate system on the sphere consisting of two orthogonal tangential unit vectors  $\mathbf{u}_i$  and  $\mathbf{v}_i$  (Fig. 2A)<sup>1</sup>. We assume that we measure the local flow component along both unit vectors subject

<sup>1</sup>For mathematical convenience, we do not take into account the hexagonal arrangement of the optical axes of the photoreceptors within the fly compound eye.

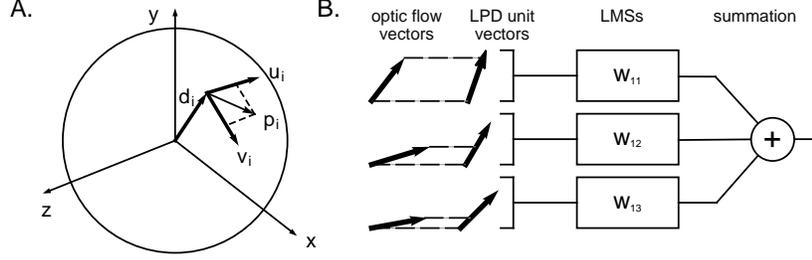


Figure 2: *A.* Sensor model: At each viewing direction  $\mathbf{d}_i$ , there are two measurements  $x_i$  and  $y_i$  of the optic flow  $\mathbf{p}_i$  along two directions  $\mathbf{u}_i$  and  $\mathbf{v}_i$  on the unit sphere. *B.* Simplified model of a tangential neuron: The optic flow and the local noise signal are projected onto a unit vector field of local preferred directions (LPDs). The projections are weighted (local motion sensitivities, LMSs) and linearly integrated. The model assumes that the integrated output encodes a egomotion component defined by either a translation or a rotation axis.

to additive noise. Formally, this means that we obtain at each viewing direction two measurements  $x_i$  and  $y_i$  along  $\mathbf{u}_i$  and  $\mathbf{v}_i$ , respectively, given by

$$x_i = \mathbf{p}_i \cdot \mathbf{u}_i + n_{x,i} \quad \text{and} \quad y_i = \mathbf{p}_i \cdot \mathbf{v}_i + n_{y,i}, \quad (1)$$

where  $n_{x,i}$  and  $n_{y,i}$  denote additive noise components with a given covariance  $C_n$  and  $\mathbf{p}_i$  the local optic flow vector. When the spherical sensor translates with  $\mathbf{T}$  while rotating with  $\mathbf{R}$  about an axis through the origin, the egomotion-induced image flow  $\mathbf{p}_i$  at  $\mathbf{d}_i$  is

$$\mathbf{p}_i = -\mu_i(\mathbf{T} - (\mathbf{T} \cdot \mathbf{d}_i)\mathbf{d}_i) - \mathbf{R} \times \mathbf{d}_i. \quad (2)$$

$\mu_i$  is the inverse distance between the origin and the object seen in direction  $\mathbf{d}_i$ , the so-called ‘‘nearness’’ (Koenderink & van Doorn, 1987). The entire collection of flow measurements  $x_i$  and  $y_i$  comprises the input to a simplified neural model which consists of a weighted sum of all local measurements (Fig. 2*B*)

$$\hat{\theta} = \sum_i^N w_{x,i} x_i + \sum_i^N w_{y,i} y_i \quad (3)$$

with local weights  $w_{x,i}$  and  $w_{y,i}$ . In this model, the local motion sensitivity is defined as  $w_i = \|(w_{x,i}, w_{y,i})^\top\|$ , the local preferred direction is parallel to the vector  $\frac{1}{w_i}(w_{x,i}, w_{y,i})^\top$ . The resulting local motion sensitivities and local preferred directions can be compared to measurements on real tangential neurons (Franz & Krapp, 2000).

As our basic hypothesis, we assume that the output of such neural models is used to estimate a egomotion component of the sensor. Since the output is a scalar, we need in the simplest case an ensemble of six neural models to encode all six rotational and translational degrees of freedom. To keep the mathematical treatment simple, we assume that the motion axes of interest are aligned with the global coordinate system. In principle, any set of linearly independent axes could be used.

The local weights of each unit are chosen to yield an optimal linear estimator for the respective egomotion component. In addition, we allow the neural models to interact linearly, such that the whole ensemble output is a linear combination of the individual neural outputs. This last step is necessary since the neural models do not react specifically to their own egomotion component due to the broad tuning of the motion detector

model (cf. Eq. (1)). The response of the neural model can be made more specific by using the output of the other neurons to suppress the signal caused by other egomotion components (Krapp, Hengstenberg, & Egelhaaf, 2001; Haag & Borst, 2003).

## 2.2 Prior knowledge

An estimator for egomotion consisting of a linear combination of flow measurements necessarily has to neglect the dependence of the optic flow on object distances. As a consequence, the estimator output will be different from scene to scene, depending on the current distance and noise characteristics. The best the estimator can do is to add up as many flow measurements as possible hoping that the individual distance deviations of the current scene from the average over all scenes will cancel each other. Clearly, viewing directions with low distance variability and small noise content should receive a higher weight in this process. In this way, prior knowledge about the distance and noise statistics of the sensor and its environment can improve the reliability of the estimate.

If the current nearness at viewing direction  $\mathbf{d}_i$  differs from the average nearness  $\bar{\mu}_i$  over all scenes by  $\Delta\mu_i$ , the measurement  $x_i$  (or  $y_i$ , respectively) can be written as (see Eqns. (1) and (2))

$$x_i = -(\bar{\mu}_i \mathbf{u}_i^\top, (\mathbf{u}_i \times \mathbf{d}_i)^\top) \begin{pmatrix} \mathbf{T} \\ \mathbf{R} \end{pmatrix} + n_{x,i} - \Delta\mu_i \mathbf{u}_i^\top \mathbf{T}, \quad (4)$$

where the last two terms vary from scene to scene, even when the sensor undergoes exactly the same egomotion.

To simplify the notation, we stack all  $2N$  measurements over the entire motion detector array in the vector  $\mathbf{x} = (x_1, y_1, x_2, y_2, \dots, x_N, y_N)^\top$ . Similarly, the egomotion components along the  $x$ -,  $y$ - and  $z$ -directions of the global coordinate system are combined in the vector  $\theta = (T_x, T_y, T_z, R_x, R_y, R_z)^\top$ , the scene-dependent terms of Eq. (4) in the  $2N$ -vector  $\mathbf{n} = (n_{x,1} - \Delta\mu_1 \mathbf{u}_1^\top \mathbf{T}, n_{y,1} - \Delta\mu_1 \mathbf{v}_1^\top \mathbf{T}, \dots)^\top$  and the scene-independent terms in the  $6 \times N$ -matrix  $F = ((-\bar{\mu}_1 \mathbf{u}_1^\top, -(\mathbf{u}_1 \times \mathbf{d}_1)^\top), (-\bar{\mu}_1 \mathbf{v}_1^\top, -(\mathbf{v}_1 \times \mathbf{d}_1)^\top), \dots)^\top$ . The entire ensemble of measurements over the sensor becomes thus

$$\mathbf{x} = F\theta + \mathbf{n}. \quad (5)$$

Assuming that  $\mathbf{T}$ ,  $n_{x,i}$ ,  $n_{y,i}$  and  $\mu_i$  are uncorrelated, the covariance matrix  $C$  of the scene-dependent measurement component  $\mathbf{n}$  is given by

$$C_{ij} = C_{n,ij} + C_{\mu,ij} \mathbf{u}_i^\top C_T \mathbf{u}_j \quad (6)$$

with  $C_n$  being the covariance of  $n$ ,  $C_\mu$  of  $\mu$  and  $C_T$  of  $\mathbf{T}$ . These three covariance matrices, together with the average nearness  $\bar{\mu}_i$ , constitute the prior knowledge required for deriving the optimal estimator.

## 2.3 Optimized linear estimator

Using the notation of Eq. (5), we write the output of the whole ensemble as a linear estimator

$$\hat{\theta} = W\mathbf{x}. \quad (7)$$

$W$  denotes a  $M \times 2N$  weight matrix where each of the  $M$  rows consists of a linear combination of the weight sets of the neural models (see Eq. (3)). The optimal weight

matrix is chosen to minimize the mean square error  $e$  of the estimator given by

$$e = E(\|\theta - \hat{\theta}\|^2) = \text{tr}[WCW^\top] \quad (8)$$

where  $E$  denotes the expectation. We additionally impose the constraint that the estimator should be unbiased for  $\mathbf{n} = 0$ , i.e.,  $\hat{\theta} = \theta$ . From Eqns. (5) and (7) we obtain the constraint equation

$$WF = \mathbf{1}_{M \times M}. \quad (9)$$

The solution minimizing the associated Euler-Lagrange functional ( $\Lambda$  is a  $M \times M$ -matrix of Lagrange multipliers)

$$J = \text{tr}[WCW^\top] + \text{tr}[\Lambda^\top (\mathbf{1}_{M \times M} - WF)] \quad (10)$$

can be found analytically and is given by

$$W = \frac{1}{2} \Lambda F^\top C^{-1} \quad (11)$$

with  $\Lambda = 2(F^\top C^{-1} F)^{-1}$ . The rows of  $F^\top C^{-1}$  correspond to the neural model of Eq. (3)<sup>2</sup>,  $\Lambda$  acts as a correction matrix that suppresses the part of the neural signal caused by the egomotion components to which the neuron is not tuned to.

When computed for the typical inter-scene covariances of a flying animal, the resulting weight sets are able to explain many of the receptive field characteristics of the tangential neurons (Franz & Krapp, 2000). However, the question remains whether the output of such an ensemble of neural models can be used for some real-world task. This is by no means evident given the fact that - in contrast to most approaches in computer vision - the distance distribution of the current scene is completely ignored by the linear estimator.

### 3 Experiments

#### 3.1 Linear estimator for an office robot

As our test scenario, we consider the situation of a mobile robot in an office environment. This scenario allows for measuring the typical motion patterns and the associated distance statistics which otherwise would be difficult to obtain for a flying agent.

The distance statistics were recorded using a rotating laser scanner. The 26 measurement points were chosen along typical trajectories of a mobile robot while wandering around and avoiding obstacles in an office environment. The recorded distance statistics therefore reflect properties both of the environment and of the specific movement patterns of the robot. From these measurements, the average nearness  $\bar{\mu}_i$  and its covariance  $C_\mu$  were computed (cf. Fig. 3, we used distance instead of nearness for easier interpretation).

The distance statistics show a pronounced anisotropy which can be attributed to three main factors: (1) Since the robot tries to turn away from the obstacles, the distance in front and behind the robot tends to be larger than on its sides (Fig. 3A). (2) The camera on the robot usually moves at a fixed height above ground (here: 0.62 m) on a flat surface. As a consequence, distance variation is particularly small at very low

<sup>2</sup>The resulting local motion sensitivities correspond exactly to those obtained from the *linear range model* in (Franz & Krapp, 2000) if one assumes a diagonal  $C$ .

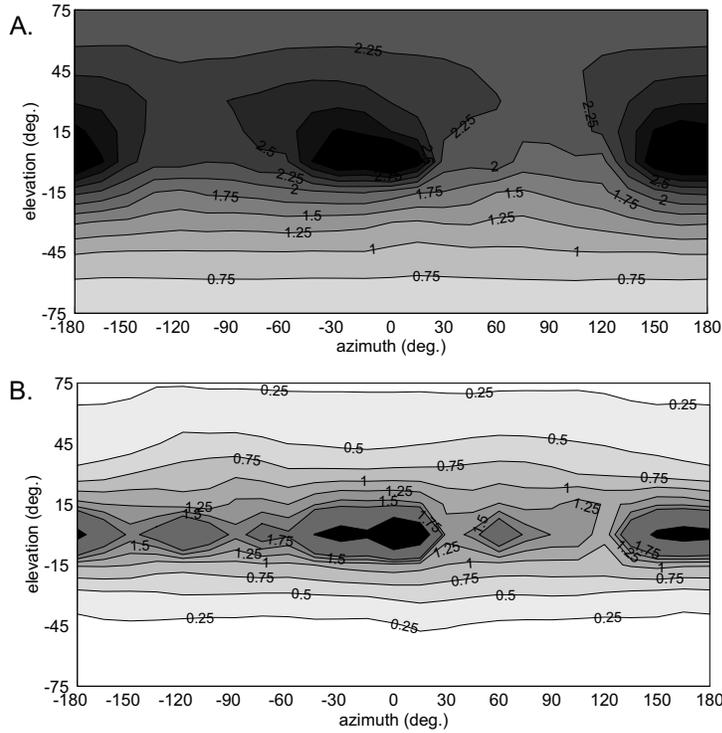


Figure 3: Distance statistics of an indoor robot (0 azimuth corresponds to forward direction; the distances on the contour lines are given in  $m$ ): *A*. Average distances from the origin in the visual field ( $N = 26$ ). Darker areas represent larger distances. *B*. Distance standard deviation in the visual field ( $N = 26$ ). Darker areas represent stronger deviations.

elevations (Fig. 3*B*). (3) The office environment also contains corridors. When the robot follows the corridor while avoiding obstacles, distance variations in the frontal region of the visual field are very large (Fig. 3*B*).

The estimation of the translation covariance  $C_T$  is straightforward since our robot can only translate in forward direction, i.e. along the  $z$ -axis.  $C_T$  is therefore 0 everywhere except the lower right diagonal entry which corresponds the square of the average forward speed of the robot (here: 0.3 m/s). The motion detector noise was assumed to be zero-mean, uncorrelated and uniform over the image, which results in a diagonal  $C_n$  with identical entries. The noise standard deviation of 0.34 deg./s was determined by presenting a series of artificially translated images of the laboratory (moving at 1.1 deg./s) to the flow algorithm used in the implementation of the estimator (see Sect. 3.2).  $\bar{\mu}$ ,  $C_\mu$ ,  $C_T$  and  $C_n$  constitute the prior knowledge necessary for computing the estimator (Eqns. (6) and (11)).

The optimal weight sets for the neural models for the six degrees of freedom (each of which corresponds to a row of  $F^T C^{-1}$ ) are shown in Fig. 4. All neural models have in common that image regions near the rotation or translation axis receive less weight. In these regions, the egomotion components to be estimated generate only small flow vectors which are easily corrupted by noise. Equation (11) predicts that the estimator will preferably sample in image regions with smaller distance variations. In

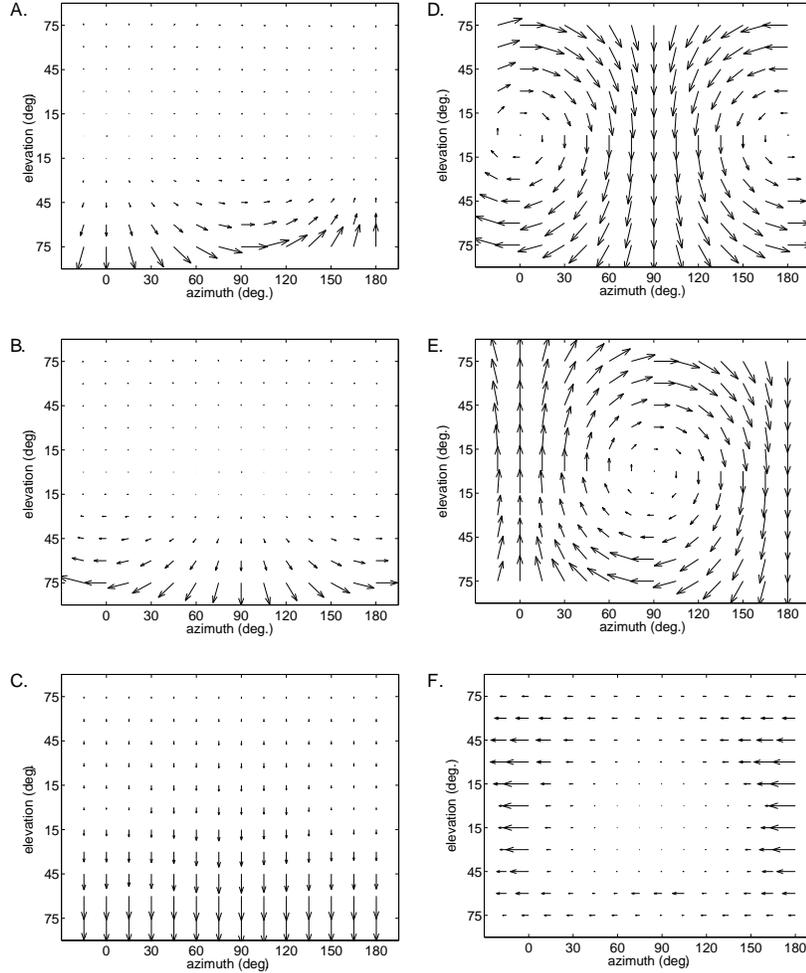


Figure 4: Neural models computed as part of the linear estimator. Notation is identical to Fig. 1. The depicted region of the visual field extends from  $-15^\circ$  to  $180^\circ$  azimuth and from  $-75^\circ$  to  $75^\circ$  elevation. The model neurons are tuned to *A.* forward translation, *B.* translations to the right, *C.* downward translation, *D.* roll rotation, *E.* pitch rotation, and *F.* yaw rotation.

our measurements, this is mainly the case at the ground around the robot (Fig. 3). The rotation-selective neural models assign higher weights to distant image regions, since distance variations at large distances have a smaller effect. In our example, distances are largest in front and behind the robot so that the neural model for yaw assigns the highest weights to these regions (Fig. 4*F*). This effect is less pronounced in the other rotational neural models because the translational flow is almost orthogonal to their local directions and thus interferes to a much lesser degree.

Although the small weights near the motion axes and the overall distribution of local directions are similar to those found in tangential neurons, our neural models show specific adaptations to the indoor robot scenario: the highly weighted ground regions are exactly the opposite to our model predictions for a flying animal where the ground region shows a stronger distance variability than regions near and above

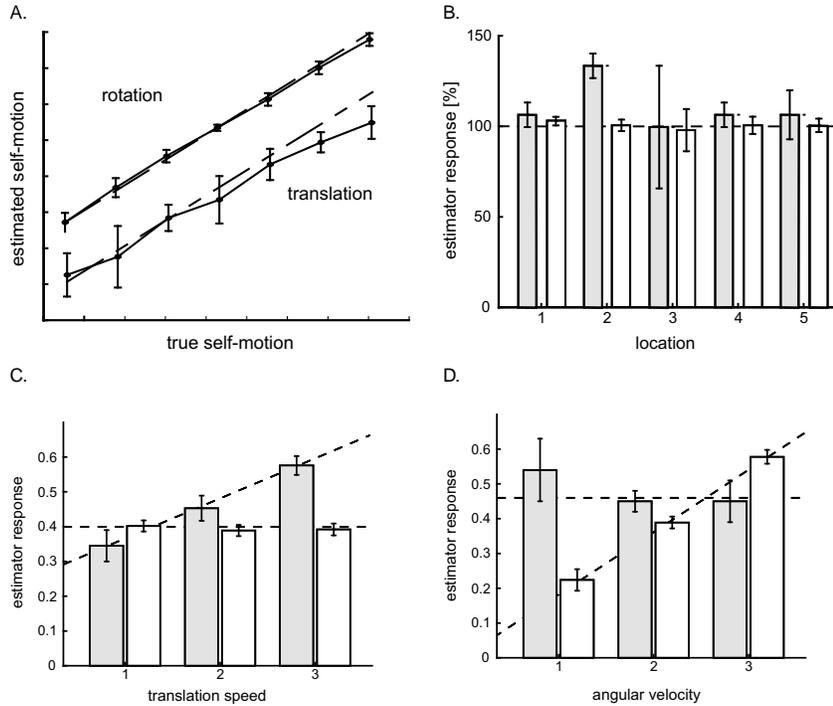


Figure 5: Gantry experiments: Results are given in arbitrary units, true rotation values are denoted by a dashed line, translation by a dash-dot line. Grey bars denote translation estimates, white bars rotation estimates. *A*. Estimated vs. real egomotion; *B*. Estimates of the same egomotion at different locations; *C*. Estimates for constant rotation and varying translation; *D*. Estimates for constant translation and varying rotation.

the horizon (Franz & Krapp, 2000). The predicted dorsoventral asymmetry with small weights in the ground region is indeed observed in the tangential neurons (see Fig. 1 and Krapp, Hengstenberg, & Hengstenberg, 1998). The strong weighting of the frontal region in the yaw neural model (Fig. 4*F*) is also corridor-specific, so it is not surprising that this feature is not present in an animal that evolved in an open outdoor environment.

### 3.2 Gantry experiments

The egomotion estimates from the ensemble of neural models were tested on a gantry with three translational and one rotational (yaw) degree of freedom. Since the gantry had a position accuracy below 1mm, the programmed position values were taken as ground truth for evaluating the estimator's accuracy.

As vision sensor, we used a camera mounted above a mirror with a circularly symmetric hyperbolic profile. This setup allowed for a 360° horizontal field of view extending from 90° below to 45° above the horizon. Such a large field of view considerably improves the estimator's performance since the individual distance deviations in the scene are more likely to be averaged out. More details about the omnidirectional camera can be found in Chahl and Srinivasan (1997). In each experiment, the camera was moved to 10 different start positions in the lab at the same height above ground

(0.62 m) as the robot camera<sup>3</sup>, but with largely varying distance distributions near and above the horizon. After recording an image of the scene at the start position, the gantry translated and rotated at various speeds and directions and took a second image. After the recorded image pairs (10 for each type of movement) were unwarped, we computed the optic flow input for the neural models using a standard gradient-based scheme (Srinivasan, 1994).

The average error of the rotation rate estimates over all trials (N=450) was 0.7°/s (5.7% rel. error, Fig. 5A), the error in the estimated translation speeds (N=420) was 8.5 mm/s (7.5% rel. error). The estimated rotation axis had an average error of magnitude 1.7°, the estimated translation direction 4.5°. The larger error of the translation estimates is mainly caused by the direct dependence of the translational flow on distance (see Eq. (2)) whereas the rotation estimates are only indirectly affected by distance errors via the current translational flow component which is largely filtered out by the local direction template. The larger sensitivity of the translation estimates to distance variations can be seen by moving the sensor at the same translation and rotation speeds in various locations. The rotation estimates remain consistent over all locations whereas the translation estimates show a higher variance and also a location-dependent bias, e.g., very close to laboratory walls (Fig. 5B). A second problem for translation estimation comes from the different properties of rotational and translational flow fields: Due to its distance dependence, the translational flow field shows a much wider range of local image velocities than a rotational flow field. The smaller translational flow vectors are often swamped by simultaneous rotation or noise, and the larger ones tend to be in the upper saturation range of the used optic flow algorithm. This can be demonstrated by simultaneously translating and rotating the sensor. Again, rotation estimates remain consistent at different translation speeds while translation estimates are strongly affected by rotation (Fig. 5C and D).

## 4 Discussion

**Egomotion estimation.** Our experiments show that it is possible to obtain useful egomotion estimates from an ensemble of linear neural models in a real-world task. Although a linear approach necessarily has to ignore the distances of the currently perceived scene, an appropriate choice of local weights and a large field of view are capable of reducing the influence of noise and distance variability on the estimates. In particular, rotation estimates were highly accurate and consistent across different scenes and different simultaneous translations. Translation estimates were not as accurate and less robust against changing scenes and simultaneous rotation. The performance difference was to be expected because of the direct distance dependence of the translational optic flow which leads to a larger variance of the estimator output. This problem can only be resolved by also estimating the distances in the current scene (e.g., in the iterative schemes in Koenderink & van Doorn, 1987; Heeger & Jepson, 1992). This, however, requires significantly more complex computations. Another reason is the limited dynamic range of the flow algorithm used in the experiments, as discussed in the previous section. One way to overcome this problem would be using an optic flow algorithm that estimates image motion on different temporal or spatial scales which is, again, computationally more expensive.

<sup>3</sup>The translational neurons in Fig. 4 for the mobile robot case assign a high weight to the ground region. As a consequence, the translation estimates strongly depend on the correct height above ground, whereas rotational neurons are only indirectly affected.

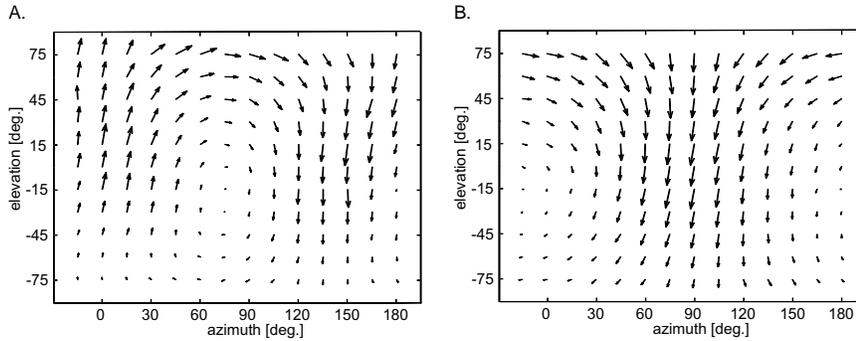


Figure 6: Hypothetical neurons constructed for *A.* roll and *B.* pitch rotation.

Our results show that the linear estimator accurately estimates rotation under general egomotion conditions and without any knowledge of the object distances of the current scene. The estimator may be used in technical applications such as image stabilization of a moving camera, or the removal of the rotational component from the currently measured optic flow. Both measures considerably simplify the estimation of distances from the remaining optic flow and the detection of independently moving objects. In addition, the simple architecture of the estimator allows for an efficient implementation at low computational costs, which are several orders of magnitude smaller than the costs of computing the entire optic flow input.

The components of the estimator are simplified neural models which - when computed for a flying animal - are able to reproduce characteristics of the tangential neuron receptive field organization, i.e. the distribution of local motion sensitivities and local preferred directions (Franz & Krapp, 2000). Our study suggests that tangential neurons may be used for self-motion estimation by linearly combining their outputs at the level of the lobula plate (e.g. Krapp et al., 2001) or at later integration stages. Evidence for the latter possibility comes from recent electrophysiological studies on motor neurons, which innervate the fly neck motor system and mediate gaze stabilization behaviour (Huston & Krapp, 2003). The possible behavioural role of such egomotion estimates, however, will critically depend on the dynamic properties of the whole sensorimotor loop, as well as on specific tuning of the motion processing stage providing the input. An example of using integrated optic flow for controlling a robotic system is described in Reiser and Dickinson (2003).

**Neural computation of egomotion.** The description of any egomotion requires at most six degrees of freedom. Therefore an ensemble of six neurons, as in our gantry experiments, would be sufficient to encode the entire egomotion of the fly. There are, however, at least 13 tangential neurons (3 HS and 10 VS neurons) in either side of the fly lobula plate which not even cover all degrees of freedom, e.g., none of the currently known receptive fields represent lift translation (reviews in Hausen, 1984, 1993; Krapp, 2000). A plausible explanation might be that the axes covered by tangential neurons - thus constituting the sensory coordinate system - are aligned with the axes used by the motor coordinate system. Recent studies on gaze stabilization in *Calliphora* suggest that in some cases the output of individual tangential neurons is connected to individual motor neurons driving certain head movements (Huston & Krapp, 2003).

Another hint comes from an interesting property of our linear model: The linearly

combined output of two model neurons corresponds to the linear combination of their respective weight sets. For the tangential neurons, this would mean that the summed output of several neurons may be treated as a superposition of their individual local response properties. The receptive fields of many tangential neurons often cover only a smaller part of the visual field, perhaps due to anatomical or developmental constraints. By summing the output of several neurons, one could build estimators with extended receptive fields covering more than one visual hemisphere. This is demonstrated in Fig. 6, where we construct a hypothetical pitch neuron from the inverted output of VS1-3 added to the output of VS8-10. Also shown in Fig. 6 is the response field of VS6 which was shown to be ideally suited to sense roll-rotations (Franz & Krapp, 2000).

**Linearity.** Finally, we have to point out a basic difference of between the proposed theory and optic flow processing in the fly: It assumes that the motion detector signals the tangential neurons integrate depend linearly on velocity (see Eq. (1)) (Reichardt, 1987). The output of fly motion detectors, however, is linear only within a limited velocity range. The motion detector output also depends on the spatial pattern properties of the visual surroundings (Borst & Egelhaaf, 1993). These properties are reflected by the tangential neurons' response properties. Beyond a certain image velocities, for instance, their response stays at a plateau when the velocity is increased. Even higher velocities result in a decrease of the neuron's response (Hausen, 1982). Within the plateau range, tangential neurons can only indicate the presence and the sign of a particular egomotion, but not the actual velocity. A detailed comparison between linear model neurons and tangential neurons shows characteristic differences. Under the conditions of the neurophysiological experiments reported in (Franz & Krapp, 2000), tangential neurons seem to operate in the plateau range rather than in the linear range. Under such response regimes a linear combination of tangential neuron outputs would not indicate the true egomotion.

Physiological mechanisms have been described, however, which may help to overcome these limitations to a certain degree. A non-linear integration of local motion detector signals, known as *dendritic gain control* (Borst, Egelhaaf, & Haag, 1995), prevents the output of the tangential neurons from saturating when its entire receptive field is stimulated. This mechanism results in a size invariant response, which still depends on velocity. Harris, O'Carroll, and Laughlin (2000) show that *contrast gain control* is of similar significance. It contributes to the neuron's adaptation to visual motion, i.e. it prevents the tangential neurons from saturating at high visual contrasts and image velocities.

Even though these mechanisms may not establish a linear dependence over the entire velocity range they may considerably extend it. Evidence supporting this idea comes from a study by Lewen, Bialek, and de Ruyter van Steveninck (2001). The authors performed electrophysiological experiments on the H1 tangential neuron in a natural outdoor environment and at bright daylight. They show that the linear dynamic range of H1 under these conditions is significantly extended compared to stimulation with a periodic grating within the same range of velocities but applied in the laboratory.

Despite these results it is still not entirely clear whether an extended linear dynamic range of the tangential neurons is sufficient to cover all needs in gaze stabilization and flight steering. Within the linear range, however, the fly might take advantage of all the beneficial properties the linear model offers. For instance, it may combine the outputs of several tangential neurons to form matched filters for particular egomotions. In case of the intrinsic tangential neuron VCH, thought to be involved in figure ground discrimination (Warzecha, Borst, & Egelhaaf, 1992), this seems to hold true. VCH re-

ceives input from several other tangential neurons the response fields of which are well characterized. By combining the response fields of the inputting tangential neurons the VCH response field is readily explained (Krapp et al., 2001). This suggests that linear combination of tangential neuron response fields may well be an option for the fly visual system to estimate egomotion.

### Acknowledgments

The authors wish to thank J. Hill, M. Hofmann and M. V. Srinivasan for their help. Financial support was provided by the Human Frontier Science Program and the Max-Planck-Gesellschaft.

### References

- Borst, A., & Egelhaaf, M. (1993). Detecting visual motion: Theory and models. In F. Miles & J. Wallman (Eds.), *Visual motion and its role in stabilization of gaze*, pp. 3 – 27 Amsterdam. Elsevier.
- Borst, A., Egelhaaf, M., & Haag, J. (1995). Mechanisms of dendritic integration underlying gain control in fly motion-sensitive interneurons. *J. Comput. Neurosci.*, **2**, 5 – 18.
- Chahl, J. S., & Srinivasan, M. V. (1997). Reflective surfaces for panoramic imaging. *Applied Optics*, **36**(31), 8275 – 8285.
- Egelhaaf, M., Kern, R., Krapp, H. G., Kretzberg, J., Kurtz, R., & Warzecha, A.-K. (2002). Neural encoding of behaviourally relevant visual-motion information in the fly. *TINS*, **25**(2), 96 – 102.
- Franz, M. O., & Chahl, J. S. (2003). Linear combinations of optic flow vectors for estimating self-motion: A real-world test of a neural model. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in Neural Information Processing Systems*, Vol. 15 Cambridge, MA. MIT Press.
- Franz, M. O., & Krapp, H. G. (2000). Wide-field, motion-sensitive neurons and matched filters for optic flow fields. *Biol. Cybern.*, **83**, 185 – 197.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Haag, J., & Borst, A. (2003). Orientation tuning of motion-sensitive neurons shaped by vertical-horizontal network interactions. *J. Comp. Physiol. A*, **189**, 363 – 370.
- Harris, R. A., O'Carroll, D. C., & Laughlin, S. B. (2000). Contrast gain reduction in fly motion adaptation. *Neuron*, **28**, 595 – 606.
- Hausen, K. (1982). Motion sensitive interneurons in the optomotor system of the fly. II. The horizontal cells: receptive field organization and response characteristics. *Biol. Cybern.*, **46**, 67 – 79.
- Hausen, K. (1984). The lobula-complex of the fly: Structure, function and significance in visual behaviour. In M. A. Ali (Ed.), *Photoreception and Vision in Invertebrates* (pp. 523 – 559). New York, London: Plenum Press.

- Hausen, K. (1993). The decoding of retinal image flow in insects. In F. A. Miles & J. Wallman (Eds.), *Visual motion and its role in the stabilization of gaze* (pp. 203 – 235). Elsevier, Amsterdam.
- Heeger, D. J., & Jepson, A. D. (1992). Subspace methods for recovering rigid motion. I.: Algorithm and implementation. *Intl. J. Computer Vision*, **7**, 95 – 117.
- Huston, S. J., & Krapp, H. G. (2003). Visual receptive field of a fly neck motor neuron. In N. Elsner & H.-U. Schnitzler (Eds.), *Göttingen Neurobiology Report 2003*, Vol. I of *Proc. 29th Göttingen Neurobiolgy Conf.*, p. 483 Stuttgart. Thieme.
- Koenderink, J. J., & van Doorn, A. J. (1987). Facts on optic flow. *Biol. Cybern.*, **56**, 247 – 254.
- Krapp, H. G. (2000). Neuronal matched filters for optic flow processing in flying insects. In M. Lappe (Ed.), *Neuronal processing of optic flow*, Vol. 44 of *Intern. Rev. Neurobiol.* (pp. 93 – 120). San Diego: Academic Press.
- Krapp, H. G., Hengstenberg, B., & Hengstenberg, R. (1998). Dendritic structure and receptive field organization of optic flow processing interneurons in the fly. *J. of Neurophysiology*, **79**, 1902 – 1917.
- Krapp, H. G., & Hengstenberg, R. (1996). Estimation of self-motion by optic flow processing in single visual interneurons. *Nature*, **384**, 463 – 466.
- Krapp, H. G., Hengstenberg, R., & Egelhaaf, M. (2001). Binocular input organization of optic flow processing interneurons in the fly visual system. *J. Neurophysiol.*, **85**, 724 – 734.
- Lewen, G. D., Bialek, W., & de Ruyter van Steveninck, R. R. (2001). Neural coding of naturalistic motion stimuli. *Network: Comput. Neural Syst.*, **12**, 317 – 329.
- Reichardt, W. (1987). Evaluation of optical motion information by movement detectors. *J. Comp. Physiol. A*, **161**, 533 – 547.
- Reiser, M. B., & Dickinson, M. H. (2003). A test bed for insect-inspired robotic control. *Phil. Trans. Royal Soc. Lond. A*, **361**, 2267 – 2285.
- Srinivasan, M. V. (1994). An image-interpolation technique for the computation of optic flow and egomotion. *Biol. Cybern.*, **71**, 401 – 415.
- Warzecha, A.-K., Borst, A., & Egelhaaf, M. (1992). Photo-ablation of single neurons in the fly visual system reveals neuronal circuit for detection of small objects. *Neurosci. Letters*, **141**, 119 – 122.