Can fly tangential neurons be used to estimate self-motion?

Matthias O. Franz[1] Titus R. Neumann[1] Michael Plagge[2]
Hanspeter A. Mallot[1] Andreas Zell[2]

[1]MPI für biologische Kybernetik, Spemannstraße 38, D-72076 Tübingen, Germany
[2]WSI, Lehrstuhl Rechnerarchitektur, Köstlinstr. 6, D-72074 Tübingen, Germany
e-mail: Matthias.Franz@tuebingen.mpg.de

Abstract

The so-called tangential neurons in the fly brain are sensitive to the typical optic flow patterns generated during self-motion. This suggests a possible involvement in the self-motion estimation process. In this study, we examine whether a simplified matched filter model of these neurons can be used to estimate self-motion from the optic flow. We present a theory for the construction of an optimal matched filter incorporating both the noise properties of the motion signal, and prior knowledge about the distance distribution of the environment. Tests on a mobile robot demonstrate that the matched filter approach works for real time camera input and the noisy motion fields computed by Reichardt motion detectors.

1 Introduction

Self-motion induces characteristic patterns of optic flow in a visual system. These patterns contain useful information about the current rotation and translation. Recent results from neurobiology provide some clues on how a biological system extracts this information from the optic flow. Krapp et al. (1998) investigated the receptive field organization of a particular class of the so-called tangential neurons (VS-neurons) in the blowfly *Calliphora vicina*. The local motion sensitivities and preferred motion directions of these neurons show a striking similarity to certain self-motion-induced flow fields (Fig. 1). Therefore, it has been argued that the tangential neurons might be involved in the extraction of self-motion parameters from the optic flow.

In this paper, we examine whether a simplified matched filter model of a tangential neuron can be used to estimate the self-motion of a mobile robot. We derive an optimal matched filter for self-motion estimation from a least square principle minimizing the variance in the filter output. The approach was originally developed to explain the observed response fields of the tangential neurons (Franz et al. 1998a). As an illustrative example, we test our theory on a mobile robot in a typical indoor environment. The experiments demonstrate that matched filters are able to extract translation and rotation velocity from optic flow under real time conditions.

2 Matched filters for optic flow patterns

2.1 A matched filter model of a tangential neuron

The local response characteristic of a tangential neuron can be modelled quite closely as the projection of the flow vector $\mathbf{p}_i$ onto a unit vector $\mathbf{u}_i$ pointing along the local preferred direction (Krapp & Hengstenberg, 1997). Assuming a linear integration over the receptive field, the excitation e of a tangential neuron in response to a motion field can be described by the simplified model (cf. Fig. 2)

$$e = \sum_i w_i(\mathbf{u}_i \cdot \mathbf{p}_i + n_i), \qquad (1)$$

where $\cdot$ denotes the dot product. The local flow measurements are taken at discrete locations with index i and projected onto a template of local preferred directions (LPDs). The local weight w_i denotes the relative local motion sensitivity (LMS), and n_i the noise in the locally measured motion signal. We assume the noise to be additive and isotropic.

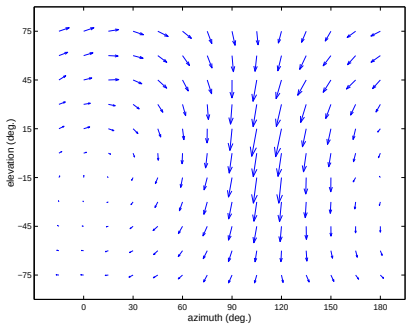

Figure 1: Mercator map of the response field of the neuron VS7. The orientation of each arrow gives the local preferred direction (LPD), and its length denotes the relative local motion sensitivity (LMS). VS7 responds maximally to rotation around an axis at an azimuth of about 30° and an elevation of about −15° (after Krapp et al., 1998).

What are the filtering properties of such a processing element? The model produces its maximal output when the flow field is parallel to the LPD template defined by the unit vector field, but also reacts to other flow fields. Similar to the matched filters in the image processing literature (Rosenfeld & Kak, 1982), the signal-to-noise ratio of the output is maximized when the stimulus field coincides with the LPD template. This is obvious in Eq. (1), since in this case the dot product is maximal in relation to the local noise signal.

2.2 *Filters for self-motion-induced flow fields*

The flow fields generated by a particular type of self-motion depend on the layout of the vision system and of the environment. Here, we consider an agent with a passive vision system that measures local flow vectors at points arranged on the unit sphere. A specific viewing direction (with index i) is described by a unit vector $\mathbf{d}_i$. When the agent translates with $\mathbf{T}$ while rotating with $\mathbf{R}$ about an axis through the origin, the self-motion-induced image flow $\mathbf{p}_i$ at $\mathbf{d}_i$ is given by (Koenderink & van Doorn, 1987)

$$\mathbf{p}_i = -\frac{(\mathbf{T} - (\mathbf{T}\cdot\mathbf{d}_i)\mathbf{d}_i)}{D_i} - \mathbf{R}\times\mathbf{d}_i, \qquad (2)$$

where D_i is the distance between the origin and the object seen in direction $\mathbf{d}_i$.

Eq. (2) shows that the local flow directions during either pure rotation or pure translation do not depend on of the object distances or on the magnitude of $\mathbf{R}$ or $\mathbf{T}$. This allows us to construct a matched filter according to our model that is tuned to the class of flow fields defined by a particular rotation or translation axis, which we refer to as the *filter axis*. To that end, the LPD template has to be parallel to the flow field class we want to detect. If we are interested in a rotatory flow field around an axis $\mathbf{a}$, the corresponding unit vector field is given by

$$\mathbf{u}_i^R = -\frac{\mathbf{R}\times\mathbf{d}_i}{\|\mathbf{R}\times\mathbf{d}_i\|} = -\frac{\mathbf{a}\times\mathbf{d}_i}{\sin\Theta_i} \qquad (3)$$

with Θ_i being the angle between viewing direction $\mathbf{d}_i$ and the filter axis $\mathbf{a}$. Analogously, the LPD template for a translation along the axis $\mathbf{a}$ is

$$\mathbf{u}_i^T = -\frac{\mathbf{d}_i\times\mathbf{a}\times\mathbf{d}_i}{\sin\Theta_i}. \qquad (4)$$

The filter signal is a linear function of the self-motion parameters $\mathbf{R}$ and $\mathbf{T}$ that can be written as the scalar product

$$e = \mathbf{R}\cdot\mathbf{C}^R - \mathbf{T}\cdot\mathbf{C}^T(D_i) \qquad (5)$$

with the vectors $\mathbf{C}^R = \sum_i w_i\mathbf{u}_i\times\mathbf{d}_i$, $\mathbf{C}^T(D_i) = \sum_i w_i\mathbf{u}_i/D_i$ and $\mathbf{u}_i$ being either $\mathbf{u}_i^R$ or $\mathbf{u}_i^T$. Thus, the filter output is proportional to the self-motion component along the filter axis, but also, to a lesser degree, to self-motion along other axes. Following Koenderink & van Doorn (1987), we call these components the *apparent* self-motion around the filter axis. Even when the self-motion parameters stay exactly the same, the filter output will vary between different trials. Apart from the noise in the motion signal, this is due to the varying distance distribution of the current scene.

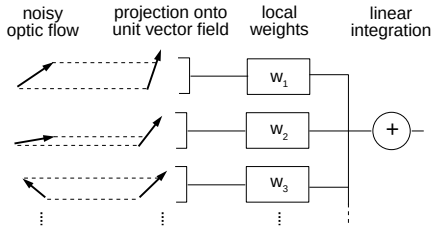

Figure 2: Matched filter model: The optic flow and the local noise signal are projected onto a unit vector field. The weighted projections are linearly integrated to give the filter output.

2.3 *Estimating self-motion using matched filters*

Eq. (5) shows that the absolute self-motion parameters cannot be recovered from the flow field alone, since the filter response depends on the unknown object distances D_i. One possibility to deal with this problem is to use prior knowledge about typical distances in the environment, e.g., by replacing the unkown term $1/D_i$ by its mean value $< 1/D_i >$. The summation over a sufficient number of local estimates will reduce the effects of individual distance deviations. In this case, the matched filter signal will be relatively independent from the particular layout of the scene as long as its statistical properties remain the same.

Consider an arrangement of three rotatory and three translatory filters tuned to different axes. The vector $\mathbf{e}$ containing the six filter outputs can be computed by the matrix product (cf. Eq. (5))

$$\mathbf{e} = \begin{pmatrix} \mathbf{C}_1^R & -\mathbf{C}_1^T \\ \vdots & \vdots \\ \mathbf{C}_6^R & -\mathbf{C}_6^T \end{pmatrix} \begin{pmatrix} \mathbf{R} \\ \mathbf{T} \end{pmatrix} \qquad (6)$$

If we replace $1/D_i$ by the mean value $< 1/D_i >$, the vectors $\mathbf{C}$ become constants and can be computed in advance. The matrix can be inverted provided that the vectors belonging to the different filters are linearly independent (which is the case in our example). Thus, the self-motion parameters can be recovered from the filter signals $\mathbf{e}$ as long as the distance statistics follow the prior assumptions. This means that a suitable arrangement of matched filters allows us to remove all apparent components from the filter output by a linear combination of the other filter signals, with the exception of the error introduced by noise and distance deviations from the mean.

3 Optimal Matched Filters

3.1 *Optimal weights for self-motion estimation*

The previous section has shown that a suitable arrangement of matched filters allows for removing all apparent components from the filter output, up to an error component due to noise and distance deviations. In this section, we derive a weight set that minimizes the error component in the filter output. The weights are chosen such that flow measurements with a high noise content and strong distance variability receive less weight.

In our model in Eq. (1), all the different noise sources are combined into a common additive noise component n_i with standard deviation Δn_i and zero mean. In addition to noise, a second error component is caused by the scattering of the object distances around their average value $\bar{D}_i$ with standard deviation ΔD_i, which results in erroneous interpretations of the underlying self-motion parameters. In order to facilitate the mathematical analysis, we have to assume that the distance variations at different points in the visual field are statistically independent. This would be ideally true in an environment consisting of small point-like objects.

Based on these assumptions, the mean square error E^2 in the filter signal can be approximated by the formula

$$E^2 \approx \sum_i w_i^2 \left(\frac{< T_i^2 > \Delta D_i^2}{\bar{D}_i^4} + \Delta n_i^2 \right) \qquad (7)$$

with $< T_i^2 >$ being the average square projection of the translation vector on the local unit vector $\mathbf{u_i}$.

In order to find a weight set minimizing the mean square error in Eq. (7), we have to impose further constraints. Here, we stipulate that the average filter signal (after removing the apparent components) should be equal to the self-motion component along its filter axis. In the case of a rotatory filter, this leads to the constraint $\sum_i w_i \sin\Theta_i = 1$. The optimal weight set can be found by solving the Euler-Lagrange equation minimizing Eq. (7) under this constraint. This leads to a simple analytic expression for the optimal weight set w_i^R of a rotatory filter

$$w_i^R = \frac{N_R \sin\Theta_i}{< T_i^2 > \Delta D_i^2 / \bar{D}_i^4 + \Delta n_i^2} \qquad (8)$$

with a suitable normalization factor N_R such that $\sum_i w_i \sin\Theta_i = 1$.

In an analogous procedure, we obtain the optimal weight set w_i^T for translation filters

$$w_i^T = \frac{N_T < 1/D_i > \sin\Theta_i}{< T_i^2 > \Delta D_i^2 / \bar{D}_i^4 + \Delta n_i^2}. \qquad (9)$$

As can be seen from Eqns. (8) and (9), the optimal solution assigns the weights accord-

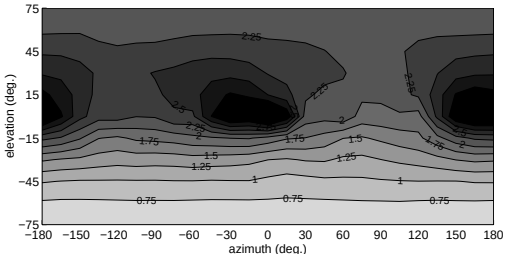

Figure 3: Average distances from the origin in the visual field ($N = 26$). Darker areas represent larger distances.

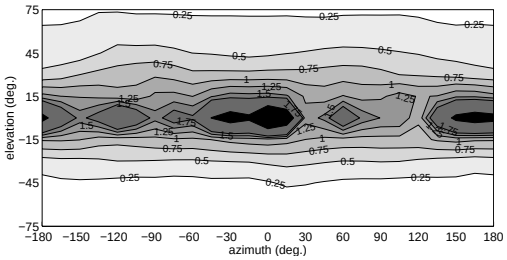

Figure 4: Distance standard deviation in the visual field ($N = 26$). Darker areas represent stronger deviations.

ing to the local variance of the corrupting factors noise and distance variability. Eqns. (8) and (9) both require prior knowledge about translation and distance statistics. In Sect. 4.1, we will show an example of these statistics obtained in an office environment.

4 Experiments

4.1 Distance statistics of office environments

The least square principle introduced above assigns a weight to each region of the visual field according to the reliability of the derived self-motion estimates. In our first experiment, we want to find out which visual region is especially suited for estimating self-motion of a mobile robot in an office environment.

The distance statistics were recorded using a laser scanner. The 26 measurement points were chosen along typical trajectories of a 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. Distance means and standard deviations are depicted in Fig. 3 and Fig. 4.

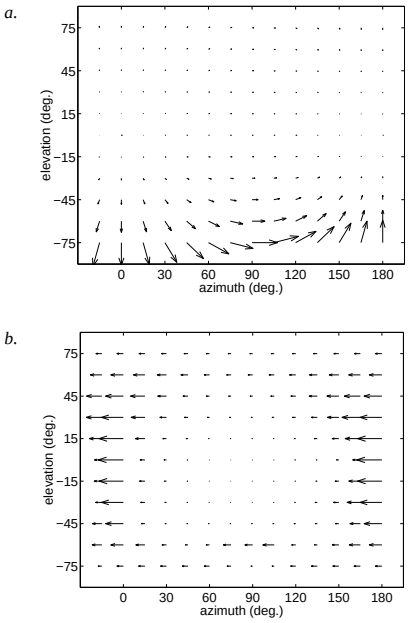

Figure 5: Optimal matched filters for office environments. The orientation of the arrows represents the local preferred direction, the absolute values of the vectors the local weights. The depicted region of the visual field extends from $-15°$ to $180°$ azimuth and from $-75°$ to $75°$ elevation. The filters are tuned to *a.* forward translation, and *b.* to rotations about the vertical axis.

The distance statistics show a pronounced anisotropy. This has three main reasons:
1. Since the robot tries to turn away from the obstacles, the free space in front and behind the robot tends to be larger than on its sides (cf. Fig. 3).
2. The camera on the robot usually moves at a fixed height above ground on a flat surface. As a consequence, distance variation is particularly small at very low elevations (cf. Fig. 4).
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 (cf. Fig. 4).
These anisotropies can be utilized in a robot vision system by preferably looking into directions with smaller relative distance variabilities.

4.2 Matched filters

The optimal filters for the two degrees of freedom of the robot (forward translation

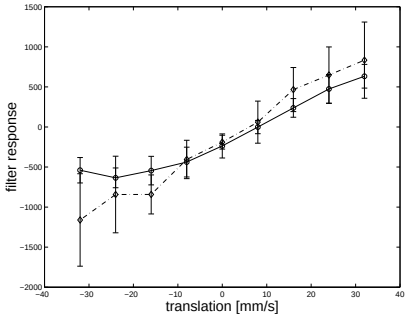

Figure 6: Translation estimates for different translation speeds (N=100). The solid line denotes estimates while keeping a straight course, the dashdot line estimates while simultaneously rotating at an angular velocity of $9°/s$.

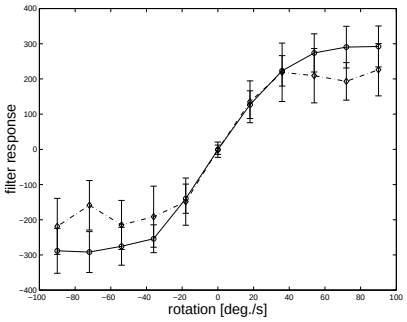

Figure 7: Rotation estimates for different angular velocities (N=100). The solid line denotes estimates while rotating on the spot, the dashdot line estimates while simultaneously translating at a speed of 24 mm/s.

and yaw rotation) for homogeneous image noise $\Delta n_i = 35\%$ in the office environment are shown in Fig. 5. The self-motion estimates are computed by reducing Eq. (6) to the two relevant dimensions. Both filters have in common that image regions near the filter axis receive less weight. In these regions, the self-motions to which the filters are tuned generate only small flow vectors which are easily corrupted by noise. Equation (9) predicts that a translation filter will preferably sample in image regions with smaller distance variations. In our measurements, this is mainly the case at the ground around the robot (cf. Fig. 4). The rotation filter weights image regions with larger distances higher, since distance variations at large distances have a smaller effect in Eq. (8). In our example, distances are largest in front and behind the robot so that the rotation filter assigns the highest weights to these regions (cf. Figs. 3).

4.3 Robot experiments

We use a modified Khepera miniature robot equipped with an omnidirectional vision system that allows for a 360° horizontal field of view extending from 10° below to 10° above the horizon. The experiments were conducted in an arena (118 × 102 cm) with model houses as visual cues. The omnidirectional camera image is sampled at 3×78 image positions. The motion field is measured using an array of Reichardt motion detectors. The resulting motion fields are very noisy and typically allow to discern 3-5 local velocities. More details about the experimental setup can be found in Franz et al. (1998b).

The signal of the translatory filter was measured for 9 translation speeds ranging from -32 to 32 mm/s. During 10 different trajectories, 10 estimates were recorded for each translation speed. The rotation filter was tested while the robot rotated at 11 angular velocities ranging from $-90°/s$ to $90°/s$. Again, 10 estimates for each angular velocity were recorded at 10 different locations.

The results are shown in Figs. 6 and 7. Both translation and rotation estimates are linear with increasing translatory and angular velocity until the saturation level of the motion detectors is reached. The average signal-to-noise ratio of the rotation estimates is smaller than that of the translation estimates. This is mainly due to the fact that the extent of the robot's visual field is more favourable to rotation estimation (cf. Sect. 4.1). In spite of the discussed problems, the presented matched filters are able to discriminate between 8-9 translation speeds and 11-12 angular velocities. The entire image processing runs at 25 Hz on a standard PC which makes the matched filters especially suitable for tasks that require fast self-motion estimates from visual input.

5 Discussion

In this paper, we presented a theory for the construction of optimal matched filters for estimating self-motion from optic flow. We derived analytical expressions for the weights from a least square principle which

minimizes the variance of the filter output. In contrast to previous approaches, prior knowledge about distance and translation statistics is incorporated into the derivation of the weights. To this end we recorded the distance statistics of a typical office environment with the help of a laser scanner. We tested the quality of the self-motion estimates generated by the matched filters on a mobile robot. The experiments demonstrated that the approach works reliably in spite of the extremely noisy motion fields computed by Reichardt motion detectors.

Elementary motion detectors. Our matched filters receive their input from correlation type EMDs which can be easily implemented in a fast computer algorithm. However, the signals of EMDs are not representing the true image velocity since their output strongly depends on the input pattern. Moreover, our EMDs are very noisy and have only a small dynamic range. All of these disadvantages could be reduced by using a more sophisticated flow algorithm, but this usually requires more computational power.

Improving the matched filter algorithm. As presented here, our approach needs distance statistics to compute the optimal filters which are hard to obtain in most environments. However, the same results can be achieved without using explicit distance measurements. Instead of indirectly estimating the variability $< T_i^2 > \Delta D_i^2 / \bar{D}_i^4$ of the translatory flow in Eqns. (8) and (9) from the distance statistics, one can also directly measure the variability of the translatory component during a long straight trajectory. The distance-induced variability can be inferred from this measurement by subtracting the noise-induced variance.

Conclusion. The present approach was originally developed to explain the observed response fields of the tangential neurons in the blowfly (Franz et al., 1998a). Our study demonstrates that a simplified model of these neurons can be used to estimate self-motion. This suggests a similar functional role of the tangential neurons in the fly brain. Moreover, the minimalistic solutions used by the insect's brain can be applied in technical systems. The computational simplicity allows for a real time implementation running at 25 Hz which makes the matched filter approach an attractive option for robot applications.

Acknowledgements

The authors wish to thank H. G. Krapp and M. Banks for helpful comments on the manuscript. Financial support was provided by the Human Frontier Science Program and the DFN.

References

Franz, M. O., Hengstenberg, R., & Krapp, H. G. (1998a). VS-neurons as matched filters for self-motion-induced optic flow fields. In N. Elsner & R. Wehner (Eds.), *New Neuroethology on the Move*, Vol. II of *Proc. 26th Göttingen Neurobiolgy Conf.*, p. 419 Stuttgart. Thieme.

Franz, M. O., Schölkopf, B., Mallot, H. A., & Bülthoff, H. H. (1998b). Learning View Graphs for Robot Navigation. *Autonomous Robots*, **5**, 111 – 125.

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

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. (1997). A fast stimulus procedure to determine local receptive field properties of motion-sensitive visual interneurons. *Vision Res.*, **37**, 225 – 234.

Rosenfeld, A., & Kak, A. C. (1982). *Digital picture processing.* London: Academic Press.