



————— Technical Report No. 61 —————

## Wide-Field, Motion-Sensitive Neurons and Optimal Matched Filters for Optic Flow

Matthias O. Franz<sup>1</sup> & Holger G. Krapp<sup>2</sup>

————— June 1998 —————

The authors wish to thank R. Hengstenberg, H. A. Mallot, M. Egelhaaf and H.-J. Dahmen for helpful comments on the manuscript. Financial support was provided by the Human Frontier Science Program and the Max-Planck-Gesellschaft.

<sup>1</sup> AG Bülhoff, E-mail: Matthias.Franz@tuebingen.mpg.de

<sup>2</sup> Lehrstuhl für Neurobiologie, Postfach 10 01 31, D-33501 Bielefeld, Germany, E-mail: Holger.Krapp@Biologie.Uni-Bielefeld.de

# Wide-Field, Motion-Sensitive Neurons and Optimal Matched Filters for Optic Flow

Matthias O. Franz & Holger G. Krapp

**Abstract.** We present a theory for the construction of an optimal matched filter for self-motion induced optic flow fields. The matched filter extracts local flow components along a set of pre-defined directions and weights them according to an optimization principle which minimizes the difference between estimated and real egomotion parameters. In contrast to previous approaches, prior knowledge about distance and translation statistics is incorporated in the form of a “world model”. Simulations indicate that the matched filter model yields reliable self-motion estimates. A comparison of the weight distribution used in the model with the local motion sensitivities of individual and small groups of interneurons in the fly visual system shows a close correspondence. This suggests that these so-called tangential neurons are tuned to optic flow fields induced by rotation or translation along a particular axis. They seem to weight the local optic flow according to the contribution of input noise and the expected variability of the translatory flow component. Their local preferred directions and motion sensitivities can be interpreted as an adaptation to the processing requirements of estimating self-motion from the optic flow.

---

## 1 Introduction

While moving through the world, biological and technical visual systems experience characteristic patterns of optic flow. These patterns are an important source of information about the animal’s self-motion parameters, i.e. the momentary translation and rotation vectors  $T$  and  $R$ , which may be useful for locomotor and gaze control. Prior to the estimation of the self-motion parameters the optic flow has to be analysed locally by elementary visual motion processing units. This bears several problems for the visual system:

1. Locally measured velocities depend on both the translatory and rotatory optic flow field. When, for instance, the rotatory component is to be measured, the superimposed translatory component may give rise to large errors (Koenderink & van Doorn, 1987).
2. Translatory flow fields depend on the distance distribution of the visible objects. Therefore, the same self-motion in different environments does not necessarily result in identical flow fields.
3. Local image motion measurements are affected by the noise in the visual input (Bouman et al., 1985) and in the synaptic signal transmission (Allen & Stevens, 1994), as

well as by the systematic errors of the elementary motion detectors due to the dependence of their output on contrast frequency, contrast and shape of the visual input (Reichardt et al., 1988).

While these problems have been treated extensively in the computer vision literature (review: Heeger & Jepson, 1992), the possible implementation of a neural system for self-motion extraction in a biological system has been elusive. In the visual system of the blowfly *Calliphora vicina*, the so-called tangential neurons were long thought to be involved in analyzing optic flow information (review: Hausen, 1984). Recently, Krapp & Hengstenberg (1996) have investigated in great detail the receptive field organization of a particular class of tangential neurons (VS-neurons) with respect to their role in the self-motion extraction process. In the present study, we try to elucidate the functional role of the VS-neurons by comparing their receptive field properties to an optimal filter for self-motion-induced flow fields.

Our approach is closely related to the concept of *matched filters* (review: Wehner, 1987). A matched filter is defined as a processing unit whose output is proportional to the cross-correlation of the filter input with a certain stimulus pattern (Rosenfeld & Kak, 1982). The local motion sensitivities and preferred directions of the tangential neurons show a striking similarity to certain

self-motion-induced flow fields. These are properties one would expect to find in a matched filter for self-motion estimation from the optic flow. Therefore, a natural question to ask is whether the matched filter concept may help to understand the receptive field organization of the tangential neurons. To that end one has to clarify how such a matched filter should be constructed, to which properties of the optic flow field it should be matched, and whether prior knowledge can help to extract relevant features for self-motion estimation. In this study, we address these questions by deriving an optimal matched filter from a least square principle. Unlike previous approaches, we explicitly incorporate prior knowledge about the distance distribution of the environment. The resulting filters do not only supply reliable self-motion estimates, e.g. for potential applications in computer vision or robotics, but also provide new insights into the functional role of the VS-neurons.

In the next section, we briefly review the experimental findings obtained in electrophysiological studies on the VS-neurons. This shows that the classical matched filter concept has to be extended to account for certain receptive field properties of the tangential neurons. In Sec. 3, we describe the self-motion estimation from the optic flow by means of a matched filter. In Sec. 4, we derive an optimal matched filter from a least square principle, and discuss the basic model assumptions. In Sec. 5, the matched filter model is checked for its performance in self-motion estimation in computer simulations and compared to the experimental results (Krapp & Hengstenberg, 1996; Krapp et al., 1998). We conclude by discussing the functional role of the VS-neurons and relating our study to other matched filter approaches.

## 2 Processing of optic flow by wide-field neurons in the fly visual system

In the third visual processing area (lobula plate) of the fly, about 60 so-called tangential neurons are known to respond directionally selective to wide-field motion stimuli (review: Hausen, 1984; Hausen & Egelhaaf, 1989). Tangential neurons integrate the signals of many retinotopically organized elementary movement detectors (EMDs; review: Reichardt, 1987) on their large dendrites. At every tiny patch in the visual field retinal image shifts are analyzed by at least 6 EMDs whose preferred directions differ according to the arrangement of neighbouring ommatidia within the

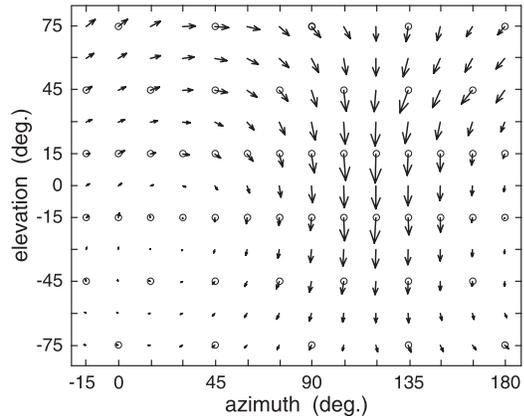


Figure 1: Mercator map of the response field of the neuron VS7. The orientation of each arrow gives the local preferred direction, and its length denotes the relative motion sensitivity. Measurement positions are marked by small circles, arrows in between are interpolated (Krapp et al., 1998).

hexagonal lattice of the fly’s compound eye (Büchner, 1976; Götz et al., 1979).

Two neuronal subsystems have been investigated more thoroughly both of which are thought to be involved in gaze- and flight stabilization: the horizontal system (HS; Hausen, 1982a,b) and the vertical system (VS; Hengstenberg et al., 1982; Hengstenberg, 1982). The 3 HS-neurons mainly integrate the output of EMDs with horizontal preferred direction. In contrast, the VS-neurons were long thought to receive first of all input from EMDs with vertical preferred direction corresponding to the strong sensitivity to vertical downward motion within their receptive fields. Some hints that a few VS-neurons also receive input from horizontally oriented EMDs (Hengstenberg, 1981) led to detailed investigation of the receptive field organization of the HS-, VS- and some other tangential neurons. The results of these investigations showed, that most of these neurons are adapted to sense particular optic flow fields rather than being sensitive to either horizontal or vertical motion only (Krapp & Hengstenberg, 1996; Krapp et al. 1998). During intracellular recordings from individual VS-neurons, the local preferred directions (LPDs) and motion sensitivities (LMSs) have been determined using a local motion stimulus (Krapp & Hengstenberg, 1997).

As an example, the distribution of LPDs and LMSs of VS7 at 48 positions within the ipsilateral hemisphere (relative to the cell body of the investigated neuron) and 4 positions within the frontal contralateral visual field is shown in Fig. 1. The

orientation of each little arrow gives the local preferred direction, and its length denotes the relative motion sensitivity. The global structure of the VS7 response field is reminiscent to an optic flow field generated by a rotation around an axis at an azimuth of about  $30^\circ$  and an elevation of about  $-15^\circ$ . There are differences between a mathematical flow field around this axis and the response field in Fig. 1: (i) The LMSs in the ventral part of the response field are smaller than the magnitude of the velocity vectors at the corresponding positions within the flow field. (ii) The LMSs around the axis of rotation are smaller than the corresponding velocity vectors. All VS-neurons have these properties in common (Krapp et al., 1998). In addition, the response fields of some VS-neurons do not comprise the entire visual hemisphere, but are confined to certain dorsal and medial regions only.

Another visual wide-field neuron in the lobula plate, the so-called Hx-neuron, neither does belong to the VS nor to the HS. The neuron was found to have a response field most similar to a translatory optic flow field. The focus of expansion of the Hx response field can be found at an azimuth of about  $135^\circ$  in the equatorial plane. Although there is an asymmetric sensitivity distribution, too, the Hx-neuron responds stronger to motion in the ventral than in the dorsal part of the visual field.

These results suggest that, in spite of their apparent similarity to optic flow fields, the LMSs of the tangential neurons cannot be understood simply as flow field templates as one would expect in a classical matched filter. The explanation of the dorsoventral asymmetry, the small LMSs near the axis of rotation or translation, and the reversed asymmetry in the Hx-neuron requires an extension of the simple template matching concept.

If we assume that the task to which these response fields are adapted is the extraction of self-motion parameters from optic flow, a model of the response fields can be derived by constructing an optimal matched filter tailor suited to solve the same problem. Thus, we may benefit in two ways: (i) by comparing such independently developed solutions, we may find a hint of how to functionally interpret the experimental results obtained in the fly, and (ii) we should end up with a model of practical significance, e.g. in robotics or computer vision.

### 3 Matched filters for optic flow patterns

#### 3.1 Extracting self-motion parameters from optic flow

We model the visual system of *Calliphora* as a collection of sensors arranged on the unit sphere, with the origin of the coordinate system centered on the head of the animal. The viewing direction of a single elementary motion detector (with index  $i$ ) is described by a unit vector  $\mathbf{d}_i$  along its axis. When the insect translates with  $\mathbf{T}$  while rotating with  $\mathbf{R}$  about an axis through the origin, the resulting image flow  $\mathbf{p}_i$  at  $\mathbf{d}_i$  is given by (Koenderink & van Doorn 1987)

$$\mathbf{p}_i = -\frac{(\mathbf{T} - (\mathbf{T} \cdot \mathbf{d}_i)\mathbf{d}_i)}{D_i} - \mathbf{R} \times \mathbf{d}_i, \quad (1)$$

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

From local measurements  $\mathbf{p}_i$  of the flow field alone, the self-motion parameters  $\mathbf{T}$  and  $\mathbf{R}$  cannot be recovered, as can be seen if one solves Eq. (1) for  $\mathbf{T}$  and  $\mathbf{R}$ :

$$\mathbf{T} = -D_i V_i^{-1} (\mathbf{p}_i + \mathbf{R} \times \mathbf{d}_i) \quad (2)$$

$$\mathbf{R} = V_i^{-1} \left( \mathbf{p}_i \times \mathbf{d}_i + \frac{1}{D_i} \mathbf{T} \times \mathbf{d}_i \right). \quad (3)$$

$V_i^{-1}$  is the inverse of the tensor  $V_i = I - \mathbf{d}_i \otimes \mathbf{d}_i$  where  $\otimes$  denotes the dyadic product and  $I$  the unit tensor. If the absolute distance  $D_i$  is unknown, the translation  $\mathbf{T}$  is only determined up to a scale factor. Moreover, the translation causes an apparent rotation  $\mathbf{A}_R = V_i^{-1} \frac{1}{D_i} \mathbf{T} \times \mathbf{d}_i$ , while the rotation causes an apparent translation  $\mathbf{A}_T = -D_i V_i^{-1} \mathbf{R} \times \mathbf{d}_i$ .

If the absolute distances of the visible objects are unknown, the actual magnitude of the translatory flow field is unpredictable, even if the momentary translation is perfectly known. As a consequence, there is no direct way of decomposing local flow measurements into their rotatory and translatory components.

A filter matched to a certain self-motion induced optic flow pattern is affected by all of these problems: it will not only react to its specific flow pattern, but also to the apparent rotation or translation caused by self-motion along other axes. Moreover, if nothing is known about the object distances, the actual contribution of the translatory flow field to the overall filter output cannot be directly determined. One way to solve these problems is to use the output of other matched filters together with plausible assumptions about the

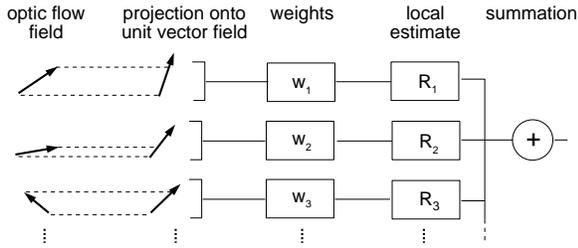


Figure 2: Matched filter model: The optic flow is projected onto a unit vector field. From the weighted projections, a local estimate of the self-motion parameter along the filter axis is computed using the factor  $1/\sin\Theta_i$  for rotatory and  $\bar{D}_i/\sin\Theta_i$  for translatory filters. Finally, all weighted local estimates are summed to give the filter output.

occurring distances to correct for the apparent rotation and translation. In the next section, we describe such a method for estimating self-motion from the output of matched filters tuned to specific optic flow fields.

### 3.2 Estimating self-motion using matched filters

The system for self-motion estimation consists of several matched filters. Each filter is tuned either to a rotatory or translatory flow field generated by self-motion along a particular axis, which we refer to as the *filter axis*. We assume that a preprocessing stage provides estimated local flow vectors as input to the matched filters. Each matched filter consists of four stages: 1. The first stage is a template of LPDs which follow the directions of a flowfield generated by a rotation or translation along the filter axis. At each position in the visual field, the input signal is obtained from the projection of the optic flow vector  $\mathbf{p}_i$  onto a unit vector  $\mathbf{u}_i$  pointing along the LPD. 2. Each projection is weighted with a local weight  $w_i$ . 3. From each projection, a local estimate of the self-motion component along the axis of the LPD template is computed. 4. In the fourth stage, all weighted local estimates are summed up to give the filter output (cf. Fig. 2). As we have pointed out in Sec. 3.1, the filter output is not only dependent on the self-motion along the filter axis, but also on the apparent rotation and translation created by self-motion along other axes. Therefore, the outputs of the other filters have to be used to remove the apparent components in the final self-motion estimate. This matched filter model is an extension of the simple template matching concept, since the local motion sensitivities (the second and third stage) do not necessarily follow the size of the flow vec-

tors generated by the particular self-motion along the filter axis.

Suppose we are interested in the rotatory flow field around a filter axis defined by the unit vector  $\mathbf{a}$ . Then the corresponding unit vector field of the filter is given by

$$\mathbf{u}_i^R = -\frac{\mathbf{a} \times \mathbf{d}_i}{\sin\Theta_i} \quad (4)$$

with  $\Theta$  being the angle between viewing direction  $\mathbf{d}_i$  and the filter axis  $\mathbf{a}$ . Analogously, the unit vector field for a translation along the axis  $\mathbf{a}$  is

$$\mathbf{u}_i^T = -\frac{\mathbf{d}_i \times \mathbf{a} \times \mathbf{d}_i}{\sin\Theta_i}. \quad (5)$$

Each projection  $m_i = \mathbf{u}_i \cdot \mathbf{p}_i$  of the actual flow on one of these unit vectors contains information about the current self-motion along the filter axis. We denote this component by the superscript  $\parallel$ , and the component orthogonal to filter axis by the superscript  $-$ . Thus, the current rotation can be decomposed in  $\mathbf{R} = \mathbf{R}^\parallel + \mathbf{R}^-$ , the translation in  $\mathbf{T} = \mathbf{T}^\parallel + \mathbf{T}^-$ .

From the projection of the local flow vector  $\mathbf{p}_i$  on  $\mathbf{u}_i^R$ , the rotation component  $R_i^\parallel = \|\mathbf{R}_i^\parallel\|$  along the filter axis can be computed using Eqns. (1) and (4) as

$$R_i^\parallel = \frac{m_i + T_i/D_i - r_i^-}{\sin\Theta_i}. \quad (6)$$

$T_i$  is the projection of the translation vector  $\mathbf{T}$  on  $\mathbf{u}_i^R$ , and  $r_i^- = \mathbf{u}_i^R \cdot \mathbf{R}^- \times \mathbf{d}_i$  denotes the projection of the rotatory flow field caused by the orthogonal component  $\mathbf{R}^-$ .  $r_i^-$  is zero if the current rotation axis exactly coincides with the filter axis  $\mathbf{a}$ . Similarly, one obtains the translation component  $T_i^\parallel = \|\mathbf{T}^\parallel\|$  along the filter axis from

$$T_i^\parallel = D_i \frac{m_i - r_i - T_i^-/D_i}{\sin\Theta_i}, \quad (7)$$

with the projections of the rotatory flow  $r_i = -\mathbf{R} \times \mathbf{d}_i \cdot \mathbf{u}_i^T$ , and of the translation component  $T_i^-$  orthogonal to the filter axis.

If the distances of the surrounding objects are unknown, the self-motion parameters cannot be computed from Eqns. (6) and (7) alone. One possibility to deal with this problem is to use prior knowledge about typical distances in the environment, e.g., by replacing the unknown distances  $D_i$  by their mean values  $\bar{D}_i$ . Provided that the current distances do not deviate too much from their mean, the summation over a sufficient number of

local estimates will then reduce the effects of individual distance deviations from the mean. In some regions of the visual field, the distance variability can be higher than in others. These regions should be weighted less in the sum. In addition, the projections are corrupted by various forms of noise and the shortcomings of the velocity detection process. We therefore assign a suitable weight  $w_i^R$  to each local estimate  $R_i^{\parallel}$  according to a least square principle which is described in Sect. 4.1.

The final rotation estimate is given by the weighted sum over the local estimates (6)

$$R_{est}^{\parallel} = \sum_i w_i^R R_i^{\parallel} \quad (8)$$

$$= \sum_i \frac{w_i^R}{\sin \Theta_i} m_i + \mathbf{T} \mathbf{C}_T - \mathbf{R}^- \mathbf{C}_R^-. \quad (9)$$

Similarly, we obtain a translation estimate from

$$T_{est}^{\parallel} = \sum_i w_i^T T_i^{\parallel} \\ = \sum_i \frac{\bar{D}_i w_i^T}{\sin \Theta_i} m_i + \mathbf{R} \mathbf{C}_R - \mathbf{T}^- \mathbf{C}_T^-. \quad (10)$$

Thus, the filter outputs  $\sum_i w_i^R m_i / \sin \Theta_i$  and  $\sum_i \bar{D}_i w_i^T m_i / \sin \Theta_i$  have to be corrected by removing the apparent rotation which arises from self-motion along other axes. The vectors  $\mathbf{C}_T = \sum_i w_i^R \mathbf{u}_i^R / \bar{D}_i \sin \Theta_i$ ,  $\mathbf{C}_T^- = \sum_i w_i^T \mathbf{u}_i^T / \sin \Theta_i$ ,  $\mathbf{C}_R = \sum_i \bar{D}_i w_i^T \mathbf{d}_i \times \mathbf{u}_i^T / \sin \Theta_i$  and  $\mathbf{C}_R^- = \sum_i w_i^R \mathbf{d}_i \times \mathbf{u}_i^R / \sin \Theta_i$  are constant and can be computed in advance, but we still need to know the values of  $\mathbf{T}$  and  $\mathbf{R}^-$ . In our system, these values are the outputs of other matched filters used as a first approximation. Three filters for each translatory and rotatory degree of freedom are sufficient to extract all self-motion parameters from optic flow. In the real organism, additional sources for self-motion estimates could be used, such as the haltere system which is capable of sensing angular velocities (Nalbach, 1994; Nalbach & Hengstenberg, 1994).

The resulting self-motion estimate can be used in turn to correct the initial estimates of the other filters. This procedure can be repeated several times until a stable solution is reached, as in the iterative algorithm of Koenderink & van Doorn (1987). However, in the kind of environments we investigated here, the first iteration was already enough to obtain very reliable estimates (cf. Sec. 5.1).

## 4 Optimal weights for self-motion estimation

### 4.1 Least square solution

The local projections  $m_i$  of the optic flow vectors onto the LPD template are likely to contain a considerable error component. Among the possible sources of error are photon noise in the visual input, synaptic transmission noise, and the characteristics of the elementary motion detectors such as their limited aperture and the dependence of their output on contrast frequency and image contrast. Although the nature of these error sources is quite different, they all lead to deviations of the measured flow component from the real one. We therefore model all the different error sources together as a common additive noise component  $n_i$  with standard deviation  $\Delta n$  and zero mean. In addition, the distances of the surrounding objects are scattered around the average value  $\bar{D}_i$  with standard deviation  $\Delta D_i$ , 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. For the derivation of the weights, we also need a probability distribution of the translation  $\mathbf{T}$  such as the example given in Sec. 4.2.

Based on these assumptions, we can ask the following question: How should the single flow measurements  $m_i$  be weighted so that the estimated self-motion parameters are as close as possible to the real ones? The answer, of course, should depend on the relative contributions of the flow signal, sensor noise and erroneous distance estimates. Starting with a matched filter for rotatory flow fields, we can formulate this question as a least square problem: Given a set of weights  $w_i$  subject to the constraint  $\sum_i w_i = 1$ , what is the optimal weight distribution that minimizes the mean squared error  $E$  of the estimated rotation  $R_{est} = \sum_i w_i R_i$  with respect to the real rotation component  $R$

$$E = \left\langle \left( \sum_i w_i R_i - R \right)^2 \right\rangle, \quad (11)$$

where the  $\langle \rangle$  denote the expectation over all trials and the  $R_i$  the local estimates from Eq. (6). Using the approximation

$$\frac{1}{D_i} \approx \frac{1}{\bar{D}_i} - \frac{D_i - \bar{D}_i}{\bar{D}_i^2} \quad (12)$$

and  $\sum_i w_i = 1$ , we obtain

$$E \approx \left\langle \left( \sum_i w_i \frac{T_i(D_i - \bar{D}_i)/\bar{D}_i^2 + n_i}{\sin \Theta_i} \right)^2 \right\rangle. \quad (13)$$

Since  $n_i$  and  $\Delta D_i$  are assumed to be statistically independent, this expression simplifies to

$$E \approx \sum_i \frac{w_i^2}{\sin^2 \Theta_i} \left( \frac{\langle T_i^2 \rangle \Delta D_i^2}{\bar{D}_i^4} + \Delta n^2 \right). \quad (14)$$

In order to find the optimal weight distribution minimizing  $E$  we have to solve the Euler-Lagrange equation

$$\frac{\partial}{\partial w_i} \left( E - \lambda \left( \sum_j w_j - 1 \right) \right) = 0, \quad (15)$$

where  $\lambda$  is a Lagrange multiplier. This finally leads to a simple analytic expression for the optimal weight distribution  $w_i^R$

$$w_i^R = \frac{N_R \sin^2 \Theta_i}{\langle T_i^2 \rangle \Delta D_i^2 / \bar{D}_i^4 + \Delta n^2} \quad (16)$$

with a suitable normalization factor  $N$  such that  $\sum_i w_i = 1$ . In an analogous procedure, one obtains the corresponding optimal weight distribution  $w_i^T$  for translation filters

$$w_i^T = \frac{N_T \sin^2 \Theta_i / \bar{D}_i^2}{\langle T_i^2 \rangle \Delta D_i^2 / \bar{D}_i^4 + \Delta n^2}. \quad (17)$$

The flow component generated by the self-motion along the filter axis follows  $\sin^2 \theta$  for rotatory filters and  $\sin^2 \theta / \bar{D}_i^2$  for translatory filters. This signal is corrupted by the noise  $\Delta n^2$  and the unpredictable variance  $\langle T_i^2 \rangle \Delta D_i^2 / \bar{D}_i^4$  of the translatory flow component due to the distance deviation from the mean. As can be seen from Eqns. (16) and (17), the optimal solution assigns the weights according to the local ratio of the signal to the variance of the corrupting factors. Eqns. (16) and (17) both require prior knowledge about translation and distance statistics. In the following section, we provide crude estimates for this kind of “world model”.

## 4.2 World model

**Distance model.** A concise statistical description of the flight patterns of *Calliphora* and its habitat is still not available in the literature, but we can make some guesses of how the probability distribution of the distances might look like. If we assume that *Calliphora* usually flies at heights

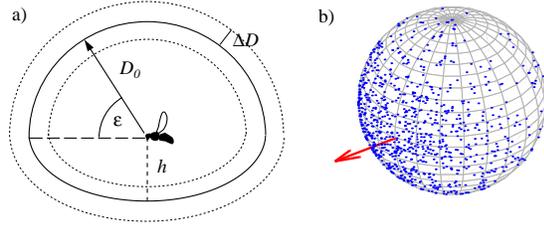


Figure 3: Simplified “world model” of *Calliphora* flying at an average height  $h$  over ground: a) Anisotropic distribution of the average distances in the visual field. The distance deviation is assumed to be independent of the viewing direction. b) 1000 samples generated by the Von Mises distribution of the translation directions. The arrow indicates the forward direction.

around 1–2 m, it is likely that, on the average, objects seen at lower elevations in the visual field will have a smaller distance than objects near or above the horizon. We model the average distance  $\bar{D}$  experienced during horizontal flight (cf. Eqns. (16) and (17)) by

$$\bar{D}(\epsilon) = \begin{cases} D_0 & : \epsilon \geq 0 \\ \frac{\beta D_0}{\sqrt{1 + (\beta^2 - 1) \cos^2 \epsilon}} & : \epsilon < 0 \end{cases}, \quad (18)$$

where  $D_0$  denotes a typical distance,  $\epsilon$  the elevation of the viewing direction and  $\beta = h/D_0$  the ratio of the average flight altitude  $h$  and  $D_0$ . The distance deviation  $\Delta D$  is chosen to be the same in all viewing directions. The resulting geometry of the distance model is that of a sphere which is flattened in the lower part of the visual field (see Fig. 3). As *Calliphora* keeps its head aligned horizontally during flight (Hengstenberg, 1993), the distance model is given in egocentric coordinates. This model clearly is a gross simplification of the real world situation, but for our purpose it is sufficient insofar that it provides a basic dorsoventral asymmetry.

**Flight directions.** To estimate the average square projection  $\langle T_i^2 \rangle$  of  $\mathbf{T}$  on  $\mathbf{u}_i^R$  in Eqns. (16) and (17), we assume a unimodal distribution for the translation direction  $\Theta$

$$p(\alpha_\Theta, \epsilon_\Theta) = N_{vM} \exp(\kappa_1 \cos \alpha_\Theta + \kappa_2 \cos \epsilon_\Theta) \quad (19)$$

around the forward direction, where  $N_{vM}$  is a normalization factor and  $\alpha_\Theta$  and  $\epsilon_\Theta$  indicate the azimuth and the elevation of the translation direction. This distribution is a three-dimensional analogue of the von Mises distribution (Batschelet, 1981).  $\kappa_1$  and  $\kappa_2$  determine the directional concentration of the distribution: smaller values lead to a broader directional distribution. In this study,

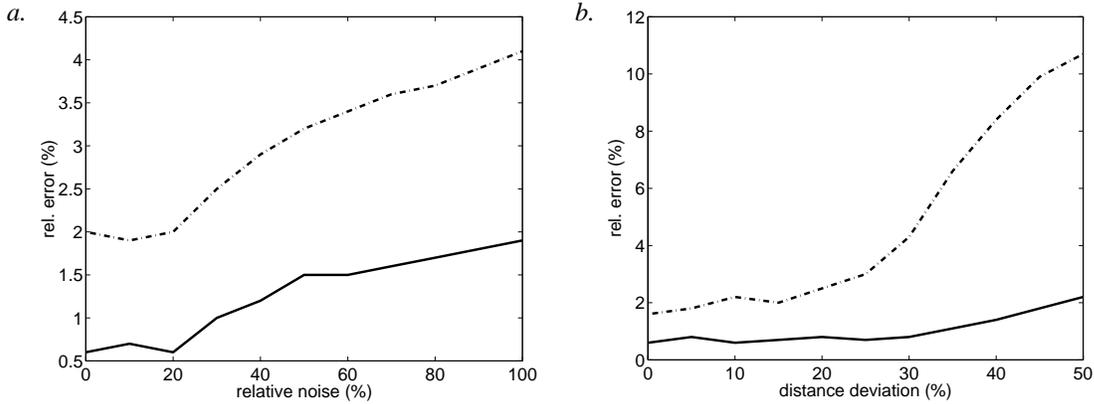


Figure 4: Relative error of self-motion estimates of 1000 samples for varying noise levels (a.) and distance deviation (b.). The solid lines denote the error of the translation estimates, dashdot lines of the rotation estimates. Noise is given relative to the average input signal, distance deviation relative to  $D_0$ .

we choose  $\kappa_1 = 2.0$  and  $\kappa_2 = 4.0$  so that the distribution becomes broader in the horizontal than in the vertical direction (see Fig. 3). Additionally, we assume that the absolute value of  $\mathbf{T}$  is distributed independently of its direction with a mean value  $T$ . The  $\langle T_i^2 \rangle$  in Eqns. (16) and (17) are then given by  $T^2 \langle p_i^2 \rangle$  with the expectation  $\langle p_i^2 \rangle$  of the square projection of the translation direction on the unit vector  $\mathbf{u}_i$ .

**Parameterization of weight distribution.** Using the normalized distance model  $D_i^n = \bar{D}_i/D_0$  from (18), a constant distance deviation  $\Delta D$  over the visual field, and the distribution of flight directions (19) with an average translation speed  $T$ , the optimal weight distributions (16) and (17) can be expressed as

$$w_i^R = \frac{N_R \sin^2 \Theta_i}{1 + \zeta \frac{\langle p_i^2 \rangle}{D_i^{n+4}}} \quad (20)$$

$$w_i^T = \frac{N_T \sin^2 \Theta_i / D_i^{n+2}}{1 + \zeta \frac{\langle p_i^2 \rangle}{D_i^{n+4}}}, \quad (21)$$

with the parameter  $\zeta = \Delta D^2 T^2 / \Delta n^2 D_0^4$  and suitable normalization factors  $N_R$  and  $N_T$ . The  $\langle p_i^2 \rangle$  are computed from 1000 direction samples generated by the von Mises distribution. Thus, the resulting weight distribution depends only on the parameters  $\beta$ ,  $\zeta$  and the position of the sensor axis.

## 5 Results

### 5.1 Testing the model

We have based our model on the assumption that the matched filters are used for extracting self-motion estimates. We first need to verify that

our model can indeed be used for this task. To that end we performed computer simulations in an artificial environment consisting of a random dot cloud. Each point of the cloud was generated randomly according to the assumed distance statistics, i.e. the average distances followed Eq. (18) with  $\beta = 0.42$ ,  $D_0 = 1.2\text{m}$ , and  $\Delta D$  was set to 20% of  $D_0$ . From the generated points, the corresponding flow vectors at the viewing directions  $\mathbf{d}_i$  were computed using Eq. (1). In each simulation run, we generated 1000 pairs of  $\mathbf{T}$  and  $\mathbf{R}$ , with  $\mathbf{T}$  chosen according to the von Mises distribution with  $T = 1.5$  m/s and  $\mathbf{R}$  uniformly distributed on the horizon with  $R = \|\mathbf{R}\| = 65^\circ/\text{s}$ , similar to the angular velocity of the test stimulus in Krapp & Hengstenberg (1997). The  $\mathbf{d}_i$  were chosen to cover the entire viewing sphere with an average mutual distance of  $2.2^\circ$ , similar to the average interommatidial angle of *Calliphora*. The resulting 9000 flow vectors were projected onto the LPD templates of the matched filters and corrupted by additive Gaussian noise with a standard deviation of 30% of the average input signal.

The weights and the LPD template were computed according to Eqns. (16) with  $\alpha = -5^\circ$ ,  $\epsilon = -4^\circ$  for a roll sensor, and  $\alpha = 73^\circ$ ,  $\epsilon = 8^\circ$  for a pitch sensor. We also tested two translatory filters (Eq. (17)), one with an axis  $\alpha = 135^\circ$ ,  $\epsilon = 0^\circ$  similar to the Hx-neuron, the other with  $\alpha = 225^\circ$ ,  $\epsilon = 0^\circ$ . The self-motion estimates were computed using Eqns. (9), and (10) from these four filters.

The relative errors  $(R - R_{est})/R_{est}$  and  $(T - T_{est})/T_{est}$  of the resulting self-motion estimates of the filter arrangement for different sensor noise levels and distance deviations are shown in Fig. 4. Due to the large number of flow measurements,

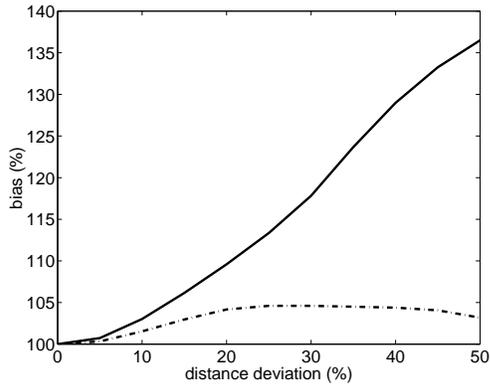


Figure 5: Bias of the estimated self-motion parameters. The solid line denotes the bias of the translation estimates, the dashdot line the bias of the rotation estimates. Distance deviation is relative to  $D_0$ .

the statistical errors caused by noise and erroneous distance estimates cancel out to a high degree. Even at a 100% noise level - i.e. only the sign of the flow component can be determined - the errors in the self-motion estimates are still below 4.2%. The higher accuracy of the translation estimates is caused by our specific choice of  $R$  and  $T$  which leads to a better signal-to-noise ratio of the translatory flow field. An interesting limitation is revealed if one looks at the bias of the estimated self-motion (Fig. 5): While the rotation estimates have almost no bias for higher distance deviations, the translation estimates show systematic errors with increasing distance deviation. This can be explained by the crude approximation in Eq. (12) which is valid only near the mean value  $\bar{D}_i$ . For higher distance deviations, the absolute values of the translatory flow vectors become asymmetrically distributed around the mean which in turn leads to systematic errors.

Nevertheless, the simulations show that the optimal filters lead to very consistent self-motion estimates even when flow vectors are very noisy and the distance deviation is very high. This consistency would also allow for correcting the systematic errors at high distance deviations by a constant factor. Clearly, the high number of flow measurements are very favourable to our approach, but our intention here is just to demonstrate its basic applicability. The effects of sparser flow fields remain a subject for further study.

## 5.2 Comparison to receptive fields of single VS-neurons

In the following sections, we investigate the relation of the proposed optimal filter model to

the measured sensitivity distributions of the VS-neurons.

From our theory, there are two possible weight distributions to which the measured data can be compared. First, we have the distribution defined in Eqns. (16) and (17). If these weights reproduce the data, we can conclude that the VS-neurons weight the flow projections  $m_i$  according to the local contributions of the rotatory flow around their preferred axis, of the input noise and of the predictability of the translatory flow component. This would mean that they act as an intermediate stage in the self-motion estimation process which extracts relevant flow patterns for subsequent processing. Alternatively, a different weight distribution is obtained if one hypothesizes that the VS-neurons extract directly rotatory velocity using a “one-shot” mechanism such as the one described by Dahmen et al. (1997). In this case, the optimal weights from Eqns. (16) and (17) have to be combined with the factors  $1/\sin\theta_i$  from Eq. (6) and  $D_i^n/\sin\theta_i$  from Eq. (7) into a common weight distribution

$$w_i^R = \frac{N_R \sin \Theta_i}{1 + \zeta_R \frac{\langle p_i^2 \rangle}{D_i^{n/4}}} \quad (22)$$

$$w_i^T = \frac{N_T \sin \Theta_i / D_i^n}{1 + \zeta_R \frac{\langle p_i^2 \rangle}{D_i^{n/4}}}, \quad (23)$$

which is proportional to  $\sin \Theta_i$  instead of  $\sin^2 \Theta_i$ .

In principle, two single neurons with different axes would be sufficient to sense any rotatory self-motions around horizontally aligned body axes. We therefore compare both weight models to the sensitivity measurements of Krapp et al. (1998) using their mean values and standard deviations. Some VS-neurons, however, are blind within large areas of the visual field (e.g. the caudolateral and caudoventral hemisphere in VS2 and VS3) which would be far from optimal if they acted as isolated sensors. A significant correspondence between measured sensitivities and the models is therefore only to be expected for VS-neurons with receptive fields covering a large part of the visual field (VS4-VS7). To test our predictions for translatory sensors, we compared our results also to the sensitivities of the Hx-neuron (Fig. 6 c, Krapp & Hengstenberg, 1996).

The theoretical distributions (16), (17), (22) and (23) have four free parameters which have to be fitted to the measured sensitivities: the azimuth  $\alpha$  and elevation  $\epsilon$  of the sensor axis, the degree of asymmetry  $\beta$ , and  $\zeta$ . The fitting was done by evaluating the  $\chi^2$  value for a given parameter

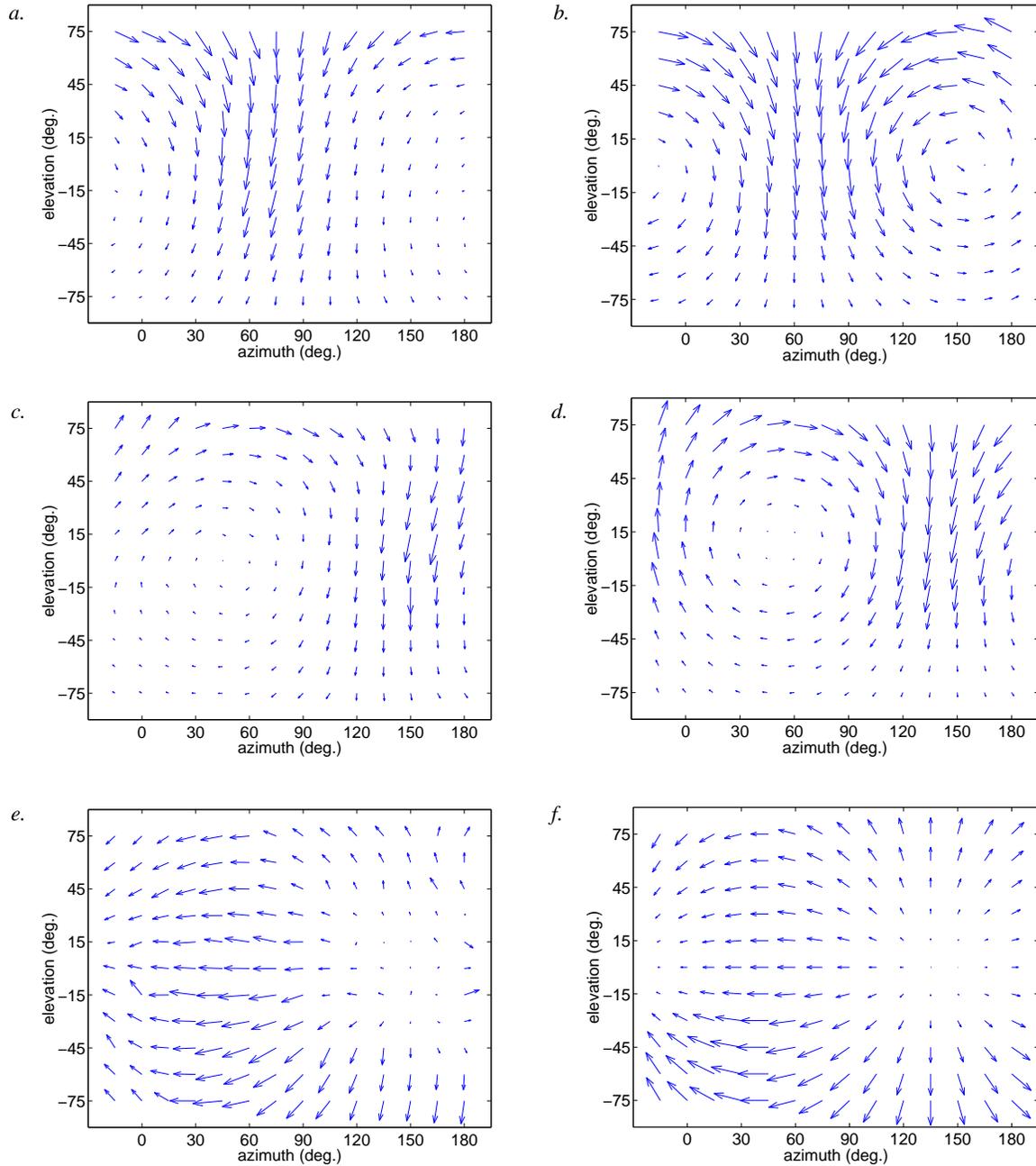


Figure 6: *a.* Averaged response field of VS4 from 5 animals. *b.* Theoretical response field obtained by fitting Eq. (18). *c.* Averaged response field of VS10 from 5 animals. This neuron is “blind” in the frontoventral area. *d.* Theoretical response field obtained by fitting Eq. (18). *e.* Measured response field of a single Hx-neuron. *f.* Theoretical response field obtained from Eq. (19) by assuming a constant standard deviation (Notation as in Fig. 1).

set using the standard deviation of the measurements. The parameter values were varied until a global minimum was reached. The step size of the parameters was 0.01 for  $\beta$ , 0.1 for  $\zeta$ , and  $1^\circ$  for the angular coordinates of the sensor axis. From the 52 measurements, those at  $-15^\circ$  elevation and  $180^\circ$  azimuth had to be discarded since due to

its body the fly could not see the visual stimulus properly. Goodness of fit was tested using a  $\chi^2$ -distribution with 47 degrees of freedom. Theoretical weight distributions with a significance  $p$  below 0.05 were rejected.

The results with  $p \geq 0.0001$  are shown in Tab. 1. Only VS4 and VS6 agreed with the weight distri-

(16)	$\chi^2$	$p$	$\beta$	$\zeta$	$\alpha$	$\epsilon$
VS4	30.9	0.97	0.41	1.0	$-26^\circ$	$-5^\circ$
VS6	59.7	0.10	0.51	2.0	$-5^\circ$	$-4^\circ$
VS8	87.5	0.0003	0.50	1.2	$35^\circ$	$-3^\circ$
VS9	90.3	0.0001	0.50	1.1	$40^\circ$	$0^\circ$

(22)	$\chi^2$	$p$	$\beta$	$\zeta$	$\alpha$	$\epsilon$
VS4	76.7	0.0004	0.40	0.9	$-21^\circ$	$-5^\circ$
VS8	85.3	0.0003	0.50	1.1	$37^\circ$	$-8^\circ$

Table 1: Results of the fitting procedure with  $p \geq 0.0001$  for single VS-neurons.  $\alpha$  and  $\epsilon$  denote azimuth and elevation of the sensor axis. The upper part of the table shows results for the weight distribution of Eq. (16), the lower part for Eq. (22). All distributions with  $p < 0.05$  are rejected.

bution described by Eq. (16). This is mainly due to the fact that these neurons do not have the blind areas in their receptive fields which are not predicted by the theory. None of the weight distributions computed according to Eq. (22) produced a fit with  $p$  higher than 0.0004 to the measured data. This suggests that – with a high probability – the VS-neurons do not produce a direct estimate of the rotatory velocity along their preferred axes. The two significant cases, VS4 (Fig. 6 *a* and *b*) and VS6, rather weight the local flow components according to Eq. (16). This weight distribution reproduces also the receptive fields of other VS-neurons outside their blind areas (cf. Fig. 6 *c* and *d*), but, as we do not have a clear description of their extent, we cannot corroborate this claim by a statistical analysis.

The qualitative features of the receptive fields of the rotatory neurons can be interpreted from Eq. (16): The retinal regions at lower elevations are less weighted due to the less predictable influence of the translatory flow field. Flow regions around the rotation axis receive less weight, too, since the rotatory flow signal is small relative to the sensor noise.

Since we presently have only one dataset of the Hx-neuron, we could not apply the same statistical procedure to compare the translatory weight sets Eq. (17) and Eq. (23) to the measured sensitivities. As a plausibility test, we assumed a constant standard deviation over the visual field and used again the same fitting procedure for Eq. (17) and Eq. (23). Both weight distributions succeeded in reproducing the elevational dependence, but the azimuthal weight distribution is again better described by the  $\sin^2 \theta$ -dependence in Eq. (17). The asymmetry of the receptive field organization is re-

versed with respect to the filters for rotatory flow fields, so that ventral retinal regions receive more weight than dorsal regions (Fig. 6 *e* and 6 *f*). This can be interpreted from our theory: In the ventral part of the visual field, the variability of the translatory flow is higher, but at the same time the signal-to-noise-ratio is better because of the larger flow vectors. In the particular type of environment we consider here, the better signal-to-noise-ratio is more important, so that the ventral regions receive higher weights.

Figure 6 demonstrates that the preferred directions of the neurons agree well with the predicted directions from Eqns. (4) and (5). It should be emphasized that the value of  $\beta$  is almost the same in all results which makes this feature of the internal world model highly reproducible. The differences in  $\zeta$  arise mainly because of the ad hoc chosen von Mises distribution which affects  $\langle p_i^2 \rangle$  in Eqns. (16) and (17). Although we did not fit the parameters  $\kappa_1$  and  $\kappa_2$  to the measurements, variations in these parameters did not lead to qualitative changes in the computed weight distributions.

Since most VS-neurons have blind areas in their receptive fields, it seems plausible that several VS-neurons will act together to form a combined sensor. Within the whole ensemble of ten different types of VS-neurons, the individual neurons could fill in the blind areas of their partners, so that the whole visual field could be sampled by such a combined sensor. In the next section, we will investigate a possible grouping of the VS-neurons for two optimal filters sensing rotations around the horizontally oriented principal body axes.

### 5.3 Comparison to ensembles of VS-neurons

As a simplified model, we assume that the output signal  $s_j$  of the VS-neuron  $j$  is the sum of the weighted projections of the local flow vectors

$$s_j = \sum_i w_i^j \mathbf{p}_i \cdot \mathbf{u}_i^j. \quad (24)$$

A sensor consisting of several neurons is then obtained by a linear combination with weights  $a_j$  of the single output signals

$$S = \sum_j a_j s_j = \sum_i \mathbf{p}_i \cdot \sum_j a_j w_i^j \mathbf{u}_i^j. \quad (25)$$

The vector sum over  $j$  can be interpreted again as a unit vector field  $\mathbf{u}_i$  together with a combined weight set  $w_i$

$$S = \sum_i w_i \mathbf{p}_i \cdot \mathbf{u}_i. \quad (26)$$

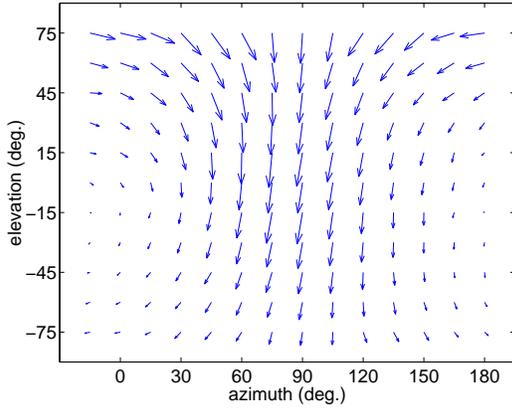


Figure 7: Averaged response field obtained by adding the local sensitivity vectors of neurons VS4, VS5, VS6 and VS7. This arrangement responds maximally to a rotatory flowfield induced by a rotation about the roll axis.

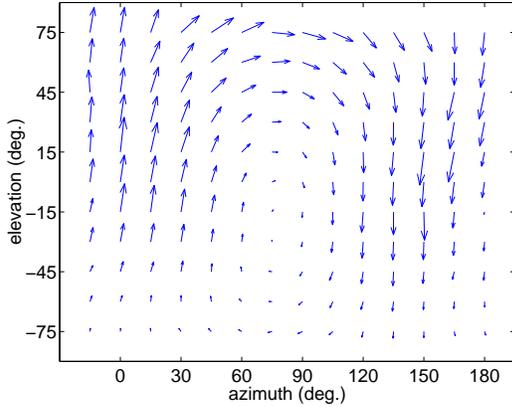


Figure 8: Averaged response field obtained by subtracting the sum of the response fields of neurons VS1 to VS3 from the sum of the response fields of VS8 to VS10 with maximal response to flowfields generated by rotations about the pitch axis.

This raises a problem: We have infinitely many possible linear combinations of the single neurons which could be compared to the theory. Fortunately, there is some evidence for a certain type of grouping. Behavioural experiments show a close correlation between the compensatory head movements of *Calliphora* in response to visual roll stimuli and the estimated summed output from VS4 to VS7 (Hengstenberg & Krapp, 1996) which makes this particular linear combination a good candidate for a roll sensor. Moreover, an orthogonal pitch sensor can be formed, if the sum of the neurons VS1 to VS3 is subtracted from the sum of VS8 to VS10. This would be the optimal counterpart to the roll sensor of VS4 to VS7 which would allow this pair of sensors to sense any rotations

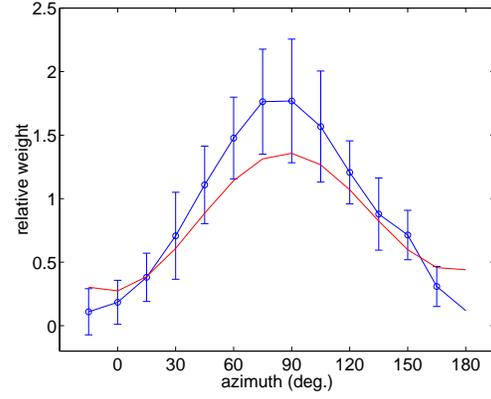


Figure 9: Horizontal slice at  $-15^\circ$  elevation through measured (circles) and predicted data (solid line) for the roll sensor.

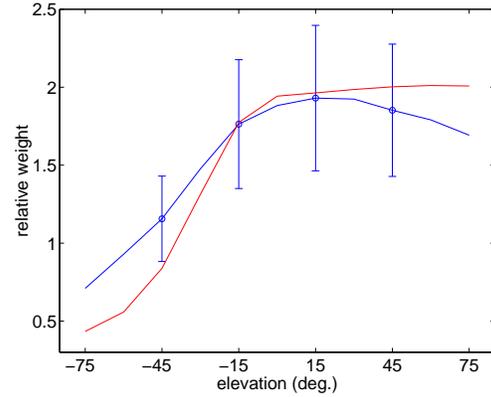


Figure 10: Vertical slice at  $75^\circ$  azimuth through measured (circles) and predicted data (solid line) for the roll sensor.

about horizontally aligned body axes. Therefore, we will apply the same fitting procedure to the hypothetical roll and pitch sensor.

(16)	$\chi^2$	$p$	$\beta$	$\zeta$	$\alpha$	$\epsilon$
roll	34.9	0.91	0.40	0.6	$-5^\circ$	$-4^\circ$
pitch	60.4	0.08	0.44	1.5	$73^\circ$	$8^\circ$

(22)	$\chi^2$	$p$	$\beta$	$\zeta$	$\alpha$	$\epsilon$
roll	338.8	0	0.50	1.5	$-8^\circ$	$-4^\circ$
pitch	66.9	0.02	0.50	1.6	$73^\circ$	$6^\circ$

Table 2: Results of the fitting procedure for the roll and pitch sensor.  $\alpha$  and  $\epsilon$  denote azimuth and elevation of the sensor axis. The upper half of the table shows results for the weight distribution of Eq. (16), the lower half for Eq. (22). All distributions with  $p < 0.05$  are rejected.

In order to construct the roll and pitch sensor

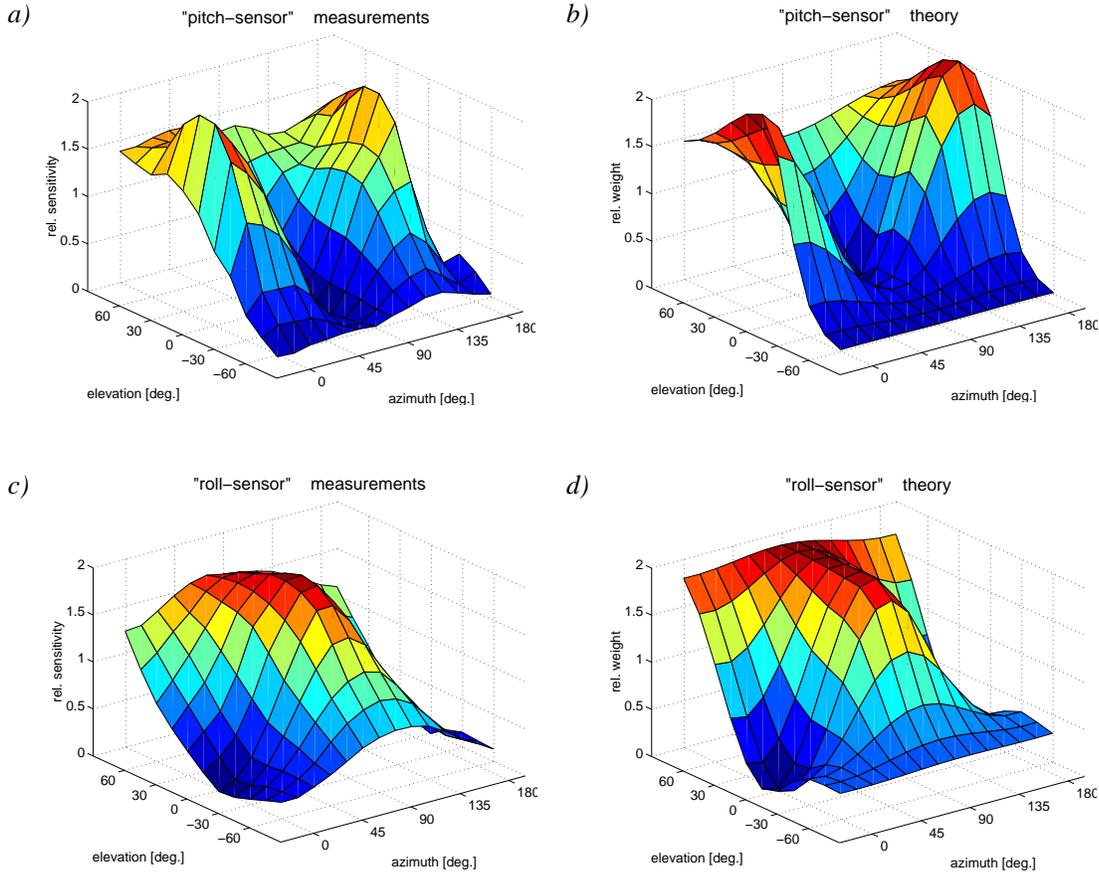


Figure 11: Weight distributions for the flow projections in the visual field: a) Measured and b) theoretically predicted weights of the pitch sensor. c) Measured and d) theoretically predicted weights of the roll sensor.

data, we used the following method: The measured sensitivities and angular orientations of 5 (3 for VS5) measurements of each single VS neuron were converted into local preference vectors. One specific sample of the hypothetical roll or pitch sensor was created by randomly choosing one of the 5 measurements for each neuron and adding or subtracting them as described above. In this way we generated 100 samples of roll and pitch sensors which were averaged to obtain the mean LMSs and LPDs. The standard deviations of the LMSs were computed from the summed variances of the single neurons. In a final step, the distributions were normalized to one. Figure 7 and 8 show the resulting averaged response fields.

We applied the same fitting procedure as above to the linearly combined data (cf. Table 2). As in the case of the single neurons, Eq. (22) is rejected, while the combined sensitivities are reproduced very well by Eq. (16) (cf. Figure 11). The blind areas of the individual neurons appear to be filled in by other neurons resulting in a relatively homoge-

neous coverage of the entire visual field. This indicates that the VS-neurons can act as an ensemble which is tuned to certain self-motion-induced rotatory flow fields. There might be other linear combinations tuned to different self-motion parameters which would allow later processing stages to use only the specific self-motion component which is relevant for its task.

The fact that Eq. (22) does not reproduce the combined data can be taken as further evidence that the VS-neurons do not directly extract self-motion estimates from the optic flow. A closer look to the measured sensitivities shows that the main difference between the tested alternative weight distributions consists in the different angular dependences: Eq. (16) predicts a  $\sin^2 \Theta$ -dependence, while Eq. (22) predicts a  $\sin \Theta$ -dependence. The horizontal slice of the roll sensor at  $-15^\circ$  (Fig. 9) illustrates that the measured weights follow very closely a  $\sin^2 \Theta$ -curve so that a  $\sin \Theta$ -curve would lead to large deviations. Most of the discrepancies between theory and data occur

at the highest and lowest elevations (cf. Fig. 10). As our theoretical world model is very crude, this could probably be corrected by using a more sophisticated world model.

## 6 Discussion

In this paper, we present a theory for the construction of optimal matched filters for estimating self-motion from optic flow. The matched filter consists of four stages: First, the flow field is projected onto a LPD template modelled as a unit vector field. Second, the single projections are weighted. Third, estimates of the self-motion parameters are extracted from the local projections and, fourth, summed over the visual field. We derive analytical expressions for the weights from an optimization principle which minimizes the difference between estimated and real egomotion parameters. In contrast to previous approaches, prior knowledge about distance and translation statistics (Fig. 3) is incorporated into the derivation of the weights. We assume an asymmetric “world model” where the average distance in the ventral visual field is smaller than in the dorsal part. We tested the quality of the self-motion estimates generated by the matched filter model in a simulated cloud of random dots. Though the self-motion estimates showed systematic errors for highly unpredictable distance statistics due to the approximation used, in general, the matched filter performance was found to be very reliable. A statistical comparison of the weight distribution used in the optimal matched filters with the measured sensitivities of individual and groups of VS-neurons shows that these neurons weight the optic flow components according to the local contributions of the input noise and of the predictability of the translatory flow component.

### 6.1 The functional role of VS-neurons

The local sensitivities of VS4 and VS6 are not statistically different from the weight distribution in Eq. (16). The other VS-neurons have more or less pronounced blind areas in their receptive fields which are not predicted by the theory. If several VS-neurons act together, each neuron could cover the blind parts of their respective partners. The analysis shows that the weights of Eq. (16) are not statistically different from two linear combinations of VS-neurons which form together a roll and a pitch sensor (cf. Fig. 11). We conclude from this result that VS-neurons can act as ensembles which are tuned to rotatory flow fields about horizontally aligned axes. They weight the single flow projec-

tions according to the contributions of the rotatory flow around their preferred axis, their noise content and the predictability of the translatory flow component. This explains the marked asymmetry of their receptive fields since the stronger relative distance variability in the ventral part of the visual field makes the translatory flow component less predictable. The asymmetry of the underlying world model is highly reproducible among the different datasets. If one looks at the receptive field of the translatory Hx-neuron (Fig. 6), one sees the reverse effect: The ventral regions receive higher weights. This can be explained by the larger translatory flow vectors in the ventral region which outweighs their unpredictability.

A second analysis showed that the hypothesis that the VS-neurons incorporate all processing stages of our matched filter can be excluded with high probability. The analysis rather suggests that the VS-neurons form only a “visual front end” to the self-motion extraction process in the sense that they weight the optic flow projections according to the information they contain about specific self-motion parameters. In other words, the VS-neurons do not directly encode rotations, but relevant aspects of the rotatory flow field. There are two functional adaptations to the processing requirements for estimating self-motion from the optic flow: first, the distribution of the LPDs building a kind of direction template and secondly, the distribution of the LMSs which weights the local flow projections according to the local signal-to-noise ratio of the self-motion parameter to which the neuron is tuned. The actual estimation of rotation and translation seemingly takes place at later processing stages of which - up to now - little is known.

In the fly’s nervous system the self-motion parameter could be represented at an intermediate processing stage which relays the output signals of the tangential neurons to the motor centers controlling the neck and flight muscles. In principle, there are at least two to three distinct ‘candidate levels’ where also the integration of the correction signals (cf. Eqns.(6) and (7)), necessary for the computation of the translation or rotation parameter (Eq. (8)), may take place: 1. At the descending neurons (DN; Strausfeld & Gronenberg, 1990; Gronenberg et al., 1995) which receive inputs from the visual system and other sensory modalities. DNs integrate the sensory signals and convey this information through the cervical connective to the motor centers in the thoracic ganglion. 2. At the level of the motor neurons driving the muscular

systems (Strausfeld et al., 1987; Gronenberg & Strausfeld, 1990). Here, the signals of the haltere system (Nalbach & Hengstenberg, 1994; Nalbach, 1994) could be exploited by means of a convergent wiring to compute the self-motion parameters. 3. Eventually, the muscles themselves could act as an integrating stage representing the projection of the momentary rotatory vector into its respective coordinate system. At the moment, we cannot answer the question in which part of the nervous system of the fly the estimation of the self-motion parameters takes place. A complete characterization of the functional role of the VS-neurons requires further investigation of how the weighted flow projections are integrated by the neurons, and how the output of these neurons is processed and integrated at later stages.

## 6.2 Model assumptions and approximations

Most of the discrepancies between theory and data can be attributed to the crude distance model. While more sophisticated world models might lead to a better correspondence, they do not necessarily provide more insights, as the asymmetry of our simple distance model already suffices to reproduce the dorsoventral asymmetry of the measured sensitivities. A more realistic world model certainly would have to be modified, e.g., by permitting a variable distance deviation over the visual field, or by including azimuthal variations of the average distances. It is quite probable, for instance, that the average frontal distances are larger than the lateral distances, since the insect, if not landing, tends to fly in the open space between obstacles. It is most obvious that the filters applied in visual environments strongly deviating from the distance model are not performing optimally any more. In other words, the recovery of the self motion parameters gets erroneous. Flies, however, are capable of flying in different environments without any observable problems. This apparent contradiction may indicate the simple fact that for flying insects the distances to visible objects in the ventral visual field are - on average - always shorter than to objects in the dorsal visual field, regardless of the environment. At least in flies, the anisotropic sensitivity distribution of the VS-neurons seems to hint in that direction.

We did not fit the von Mises distribution of the translation directions to the measured data. This primarily affects the parameter  $\zeta$  which is directly related to the translatory flow component (cf. Eqns. (20) and (21)). An appropriate choice of

the translation statistics might lead to more consistent values of  $\zeta$  among the different datasets, but this would require a concise statistical description of *Calliphora's* flight trajectories which is currently not available in the literature.

To compare the weight distributions obtained from the matched filter model with the hypothetical roll and pitch sensors, we linearly summed up the weighted inputs of some VS-neurons. Neuronal elements usually do not sum up their inputs ad infinitum but saturate at a level given by a maximum spike rate or a maximum depolarisation of the membrane. This has been shown for the tangential neurons, for instance (Hengstenberg, 1982; Haag et al., 1992).

The linear approximation of the distance dependence in Eq. (12) is only valid for smaller distance deviations. While this does not result in significant discrepancies between theory and measured sensitivities, the approximation leads to systematic errors of some filters in simulated random dot clouds with high distance variability. As the self-motion estimates are also consistent at high distance deviations, these systematic errors could be compensated by a constant factor, or by higher order correction terms in Eq. (12).

Our matched filter model assumes that the direction and velocity of local image shifts are evaluated. The output of EMDs is clearly not representing the true velocity of the image shift. In contrast, the output of correlation type EMDs depends on the contrast frequency of the input pattern; i.e. the ratio of the angular displacement over the spatial frequency of the pattern. In addition, the response depends on the contrast of the visual input pattern (Reichardt, 1987). The simulation results for 100% noise (cf. section 4.1) indicate that the matched filter model can tolerate a considerable degree of corruption in the input, but it needs to be tested in further study whether they are sufficiently robust for correlation type EMD input. Another problem which is always inherent to local direction and velocity measurements, the aperture problem, has to be taken into account, too. The aperture problem can be overcome - at least partially - by spatially integrating over many local measurements covering an extended part of the visual field.

So far, we have not studied the dynamical behaviour of our model. In the real world, the visual input and the EMD output may change rapidly during locomotion which might lead to an unstable filter output. Although our simulations suggest that the filter output does not vary strongly

between different trials, this still needs to be verified for more realistic visual input.

### 6.3 Relation to other matched filter approaches

*Dahmen et al. (1997)* derived a matched filter as a special case of the iterative algorithm by *Koenderink & van Doorn (1987)*. Similar to our approach, this algorithm is derived from a least square principle. Instead of minimizing the difference between actual and estimated self-motion parameters, *Koenderink & van Doorn* minimize the difference of the measured flow field to a flow field derived from the self-motion estimates. In contrast to our approach, they assume no prior knowledge about distance statistics. Thus, in addition to the self-motion parameters, also distance estimates have to be extracted from the optic flow. Therefore, their algorithm (and also those discussed below) allows only to estimate translation direction, not its absolute value.

Assuming a spherical environment, *Dahmen et al. (1997)* showed that the first iteration of the algorithm of *Koenderink & van Doorn (1987)* can be implemented by a matched filter similar to ours. *Dahmen et al. (1997)* tested their filter using simulated noisy flow as input. They report excellent performance in spherical environments, in many cases close to that of the iterative algorithm. The weights of their matched filter can also be derived from Eqns. (16) and (17) by assuming constant average distance, distance deviation and noise over the visual field. This suggests that our least square principle can be understood as an extension of the approach of *Koenderink & van Doorn (1987)* to cases, where prior knowledge about distance statistics is available.

*Perrone (1992)* presented a matched filter model of self-motion estimation in the primate visual cortex. In his approach, the motion field was sampled at each image position by several sensors tuned to different velocity vectors. The matched filters use only those velocity signals as input which are consistent with a given set of self-motion parameters. The self-motion estimate is derived from the most active filter using a winner-takes-all strategy. Every combination of self-motion parameters needs its own matched filter. Although this approach requires a huge number of filters for the general case, it has the advantage that no correction terms for apparent rotations and translation are needed. The number of filters can be somewhat reduced by constructing only filters for the most probable parameter sets, e.g. by consider-

ing only fixating eye movements (*Perrone & Stone, 1994*). In contrast, our approach requires only six filters, one for each translatory and rotatory degree of freedom, to determine all self-motion parameters. Any additional matched filter provides redundant information which probably could be used to improve the robustness and accuracy of the self-motion estimates. This appears to be also a strategy employed by the fly visual system with its 60 tangential neurons, but the possible advantages of such a redundant sampling have to be clarified in further study.

*Fermüller & Aloimonos (1995)* use the fact that every set of self-motion parameters induces certain patterns in the flow directions. These patterns are used as filters which are matched to the measured flow components along pre-defined curves in the image. Similar to *Perrone (1992)*, every combination of self-motion parameters requires its own motion template. Since *Fermüller & Aloimonos (1995)* discard part of the available information in the flow, they have to provide an extensive decision procedure to arrive at a consistent self-motion estimate. They use only the sign of selected flow components as input, since the limited aperture of local motion detectors allows only for detecting the projection of the optic flow on the local image gradient (aperture problem, cf. *Horn & Schunk, 1981*). This problem comes into play when the input consists of grey value image sequences instead of random dot patterns as in our simulations. As we have not tested our matched filter model on real image sequences, it is currently unclear how they would perform in real world situations. Nevertheless, the results in Fig. 4 for 100% noise in the EMD input show that our filters work also when only the sign of the optic flow projection can be determined.

### 6.4 Conclusion

We derived a matched filter model from a theoretical optimality criterion. As we have shown, these optimal filters provide some hints on the functional role of certain tangential neurons in the fly visual system. Their performance in simulations makes them also potentially useful for technical applications in robots or computer vision systems, since they require only small computational resources. Before this can be realized, some open questions remain to be investigated, e.g., how a set of matched filters should be coupled to arrive at stable self-motion estimates, or how much a redundant sampling with several filters can improve the reliability of self-motion estimates. The an-

swers to these questions can be fruitful not just for biological research, but also for technical applications which could be considered the real world test of the theory we presented here.

## References

- Allen, C., & Stevens, C. F. (1994). An evaluation of causes for unreliability of synaptic transmission. *Proc. Natl. Acad. Sci. USA*, **91**, 10380–10383.
- Batschelet, E. (1981). *Circular statistics in biology*. London: Academic Press.
- Bouman, M. A., van Grind, W. A., & Zuidema, P. (1985). Quantum fluctuations in vision. In E. Wolf (Ed.), *Progress in Optics*, Vol. XXII, pp. 79 – 144. North Holland, Amsterdam, New York, Oxford, Tokyo.
- Buchner, E. (1976). Elementary movement detectors in an insect visual system. *Biol. Cybern.*, **24**, 85 – 101.
- Dahmen, H., Wüst, R. W., & Zeil, J. (1997). Extracting egomotion parameters from optic flow: principal limits for animals and machines. In M. V. Srinivasan & S. Venkatesh (Eds.), *From living eyes to seeing machines*, pp. 174 – 198. Oxford University Press, Oxford, New York.
- Fermüller, C., & Aloimonos, Y. (1995). Qualitative egomotion. *Intl. J. Computer Vision*, **15**, 7 – 29.
- Götz, K. G., Hengstenberg, B., & Biesinger, R. (1979). Optomotor control of wing beat and body posture in *Drosophila*. *Biol. Cybern.*, **35**, 101 – 112.
- Gronenberg, W., Milde, J. J., & Strausfeld, N. J. (1995). Oculomotor control in calliphorid flies - organization of descending neurons to neck motor-neurons responding to visual-stimuli. *J. Comp. Neurol.*, **361**, 267 – 284.
- Gronenberg, W., & Strausfeld, N. J. (1990). Descending neurons supplying the neck and flight motor of diptera: Physiological and anatomical characteristics. *J. Comp. Neurol.*, **302**, 973 – 991.
- Haag, J., Egelhaaf, M., & Borst, A. (1992). Dendritic integration of motion information in visual interneurons of the blowfly. *Neurosci. Lett.*, **140**, 173 – 176.
- Hausen, K. (1982a). Motion sensitive interneuron in the optomotor system of the fly. II. The horizontal cells: Receptive field organization and response characteristics. *Biol. Cybern.*, **46**, 67 – 79.
- Hausen, K. (1982b). Motion sensitive interneurons in the optomotor system of the fly. I. The Horizontal Cells: Structure and signals. *Biol. Cybern.*, **45**, 143 – 156.
- 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. Plenum Press, New York, London.
- Hausen, K., & Egelhaaf, M. (1989). Neural mechanisms of visual course control in insects. In R. C. H. D. G. Stavenga (Ed.), *Facets of vision*, pp. 391 – 424. Springer, Berlin, Heidelberg.
- Heeger, D. J., & Jepson, A. D. (1992). Subspace methods for recovering rigid motion. I: Algorithm and implementation. *Intl. J. Computer Vision*, **7**, 95 – 117.
- Hengstenberg, R. (1981). Rotatory visual responses of vertical cells in the lobula plate of Calliphora. *Verh. Dtsch. Zool. Ges.*, **74**, 180.
- Hengstenberg, R. (1982). Common visual response properties of giant vertical cells in the lobula plate of the blowfly calliphora. *J. Comp. Physiol. A*, **149**, 179 – 193.
- Hengstenberg, R. (1993). Multisensory control in insect oculomotor systems. In F. Miles & J. Wallman (Eds.), *Visual motion and its role in stabilization of gaze*, pp. 285 – 298. Elsevier, Amsterdam, Londo, New York, Tokyo.
- Hengstenberg, R., Hausen, K., & Hengstenberg, B. (1982). The number and structure of giant vertical cells (VS) in the lobula plate of the blowfly Calliphora erythrocephala. *J. Comp. Physiol. A*, **149**, 163 – 177.
- Hengstenberg, R., & Krapp, H. G. (1996). Distribution of roll motion sensitivity in the eyes of Calliphora: a comparison between neurons and behaviour. In H.-U. S. N. Elsner (Ed.), *Proc. 24th Göttingen Neurobiology Conf.*, Vol. I, p. 349. Thieme, Stuttgart.
- Horn, B. K. B., & Schunk, B. G. (1981). Determining optical flow. *Artificial Intelligence*, **17**, 185 – 203.

- 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. (1996). Estimation of self-motion by optic flow processing in single visual interneurons. *Nature*, **384**, 463 – 466.
- 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.
- Nalbach, G. (1994). Extremely non orthogonal axes in a sense organ for rotation: Behavioural analysis of the dipteran haltere system. *Neuroscience*, **61**, 149 – 163.
- Nalbach, G., & Hengstenberg, R. (1994). The halteres of the blowfly *Calliphora*. II. Three-dimensional organization of compensatory reactions to real and simulated rotations. *J. Comp. Physiol. A*, **175**, 695 – 708.
- Perrone, J. A. (1992). Model for the computation of self-motion in biological systems. *J. Opt. Soc. Am. A*, **9**, No. 2, 177 – 194.
- Perrone, J. A., & Stone, L. S. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision Res.*, **34**, 2917 – 2938.
- Reichardt, W. (1987). Evaluation of optical motion information by movement detectors. *J. Comp. Physiol. A*, **161**, 533 – 547.
- Reichardt, W., Egelhaaf, M., & R. W. Schloegel, . (1988). Movement detectors provide sufficient information for local computation of 2-D velocity field. *Naturw.*, **75**, 313 – 315.
- Rosenfeld, A., & Kak, A. C. (1982). *Digital picture processing*. London: Academic Press.
- Strausfeld, N. J., & Gronenberg, W. (1990). Descending neurons supplying the neck and flight motor of diptera: Organization and neuroanatomical relationships with visual pathways. *J. Comp. Neurol.*, **302**, 954 – 972.
- Strausfeld, N. J., Seyan, H. S., & Milde, J. J. (1987). The neck motor system of the fly *Calliphora erythrocephala*. I. Muscles and motor neurons. *J. Comp. Physiol. A*, **160**, 205 – 224.
- Wehner, R. (1987). Matched filters - neuronal models of the external world. *J. Comp. Physiol. A*, **161**, 511 – 531.