Network: Computation in Neural Systems
December 2007; 18(4): 343–374

informa
healthcare

Local statistics of retinal optic flow for self-motion through natural sceneries

DIRK CALOW & MARKUS LAPPE

Department of Psychology, Westf.- Wilhelms University, Fliednerstr. 21, 48149 Münster, Germany

(Received 23 February 2007; revised 9 July 2007; accepted 21 August 2007)

Abstract
Image analysis in the visual system is well adapted to the statistics of natural scenes. Investigations of natural image statistics have so far mainly focused on static features. The present study is dedicated to the measurement and the analysis of the statistics of optic flow generated on the retina during locomotion through natural environments. Natural locomotion includes bouncing and swaying of the head and eye movement reflexes that stabilize gaze onto interesting objects in the scene while walking. We investigate the dependencies of the local statistics of optic flow on the depth structure of the natural environment and on the ego-motion parameters. To measure these dependencies we estimate the mutual information between correlated data sets. We analyze the results with respect to the variation of the dependencies over the visual field, since the visual motions in the optic flow vary depending on visual field position. We find that retinal flow direction and retinal speed show only minor statistical interdependencies. Retinal speed is statistically tightly connected to the depth structure of the scene. Retinal flow direction is statistically mostly driven by the relation between the direction of gaze and the direction of ego-motion. These dependencies differ at different visual field positions such that certain areas of the visual field provide more information about ego-motion and other areas provide more information about depth. The statistical properties of natural optic flow may be used to tune the performance of artificial vision systems based on human imitating behavior, and may be useful for analyzing properties of natural vision systems.

Keywords: *Optic flow, natural ego-motion, statistics, entropy estimation, mutual information estimation*

Correspondence: Dr Dirk Calow, Department of psychology, Westf.-Wilhelms University, Fliednerstrasse 21, 48149 Münster, Germany. E-mail: calow@psy.uni-muenster.de

ISSN 0954-898X print/ISSN 1361-6536 online/07/040343–374 © 2007 Informa UK Ltd.
DOI: 10.1080/09548980701642277

Introduction

Often the brain has to analyze sensory signals which are ambiguous. Ambiguity arises from the spatial and/or temporal properties of the perceptual sensors, from noise introduced by the perceptual sensors, and from noise created by the environment. To (re)construct perception, the brain may use statistically plausible predictions and/or statistical models of the signal-sending environment. The resources of a signal processing system are usually limited, and therefore the range of signals that can be processed is bounded. Nonlinear processing schemes that include knowledge of the statistics of the signals can enable the system to be more sensitive for signals which occur very frequently, and to attach less value to signals which are very unlikely to occur. Such statistically efficient processing schemes restrict the limited resources of the system to the range of statistically probable signals. Therefore, evolutionary adaptations of the perceptual areas of the brain to the statistics of natural environments, are plausible. Effects of such adaptations are seen in gestalt laws (Elder and Goldberg 2002; Krüger and Wörgötter 2002) and in efficient encoding schemes (Barlow 1961; Lauglin 1981).

In the visual modality, several researchers invested effort to reveal the statistics of natural environments, and to link it with the neural representation of the sceneries (Laughlin 1981; Atick and Redlich 1992; Rudermann and Bialek 1994; Olshausen and Field 1996; Krüger 1998; van Hateren and Rudermann 1998; Zetsche and Krieger 2001; Simon celli and Olshausen 2001; Berkes and Wiskott 2002; Betsch et al. 2004). Their investigations are largely restricted to static attributes of natural scenes, however, even when dynamic stimulus material was used (van Hateren and Rudermann 1998; Betsch et al. 2004). Furthermore, the resulting statistics are treated as independent of the position in the field of view. The properties of motion signals elicited on optic detectors by ego-motion within natural sceneries strongly depend on the position in the view field (Zanker and Zeil 2005). To investigate the statistics of these motion signals therefore requires an analysis of distributions of flow vectors with respect to their visual field position.

Optic flow generated by self motion encodes much information about the direction of ego-motion, the velocity, the distances of potential obstacles, and the structure of the environment (Gibson 1950, 1966). Animals use this information for path planning, obstacle avoidance, ego-motion control, and foreground–background segregation (see Lappe (2000b) for an overview). The motion signals of the optic flow are processed in specialized motion-processing brain areas (Saito et al. 1986; Albright 1989; Duffy and Wurtz 1991; Lappe et al. 1996). It is likely that the motion-processing pathway of the brain uses statistical properties of natural flow fields to efficiently encode natural optic flow, and to reconstruct the true motion field from the motion signals in early motion detectors. We hypothesize that the brain has involved mechanisms of extracting information from optic flow which benefit from statistical dependencies of the elicited optic flow on the properties of the natural environment and natural motion situations. An investigation of the local statistical properties of optic flow can be the starting point to reveal such connections.

The analysis of the statistics of optic flow may be undertaken on the true motion signals (Ivins et al. 1999; Calow et al. 2004; Roth and Black 2005), or on the signals

obtained from early motion detectors (Fermuller et al. 2001; Kalkan et al. 2005; Zanker and Zeil 2005). The latter approaches analyze the combination of properties of the motion field generated by ego-motion with the properties of particular motion detectors. Since we are interested in the statistical properties of the motion field itself we need to analyze the true motion signals. Therefore, we need a large number of true motion fields generated by natural ego-motion through a natural environment.

A method to collect a sufficient number of true optic flow fields was introduced in Calow et al. (2004). Based on the Brown range image database (Huang et al. 2000) true motion fields were generated by biologically plausible ego-motion and first results of the first-order statistics of retinal optic flow fields were reported. Roth and Black (2005) used this method to investigate the statistics of optic flow and elementary optic flow components. Since their work mainly focused on aspects of machine vision, the ego-motion parameters underlying the optic flow fields were obtained from ego-motions of hand-held or car-mounted cameras. The resulting statistics were treated as independent of the position in the field of view. Our investigation is dedicated to the local statistics of the true retinal motion signals occurring in biological vision during natural human-like ego-motion. In natural locomotion, eye movement reflexes stabilize gaze on objects of interest in the scene (Solomon and Cohen 1992; Lappe et al. 1997; Niemann et al. 1999) such that natural ego-motion is always a combination of body movement and eye movement. The combination of body movement, eye movement, and depth structure of the visual environment determines the structure of the optic flow on the retina (Lappe et al. 1999). Our investigation of the statistical properties of the flow field is therefore based on a combination of walking and eye-movement reflexes.

We use the term local statistics to note the statistical properties of the distributions of retinal velocities and their statistical dependencies on depth and ego-motion for certain positions in the field of view. The correlations between motion signals of different positions are not part of our notion of local statistics.

We see the purpose of our study in providing basic information and quantitative data on the statistics of retinal motion signals. This information can be used to predict sensitivity ranges of neurons in the motion-processing pathway of the brain. Future work will focus on the examination of the hypothesis that these neurons efficiently encode distributions of naturally occurring retinal motion signals. The knowledge of the statistics is crucial for that purpose. Furthermore, the knowledge of the statistical properties of retinal motion signal is an important tool in creating experimental paradigms that focus on natural motion stimuli. Comparisons between natural and unnatural motion situations are necessary to reveal how the motion-processing pathway is adapted to the statistics of the natural environment. Our investigation can also provide prior knowledge for creating probabilistic models of the motion–processing pathway of the brain based on Bayesian inference (Weiss and Fleet 2001).

Since the local statistics of optic flow are tightly linked to the statistics of the depth structure of natural scenes and to the statistics of the ego-motion parameters, the information about the depth map of the current scene and the ego-motion situation is encoded in the retinal flow. However, the generation of optic flow maps from a 5-dimensional parameter space (walking speed, heading, depth, and depth of the

fixation point) to a 2-dimensional flow vector (cf. Equations 5 and 6). Therefore, the information about the underlying parameters is condensed in the flow vector and cannot be extracted from an individual flow vector directly. Recovery of heading, for instance, requires the combined information from several flow vectors (Longuet-Higgins and Prazdny 1980). However, different areas in the field of view show different statistical dependencies of the components of the optic flow on heading and depth. By focusing the analysis on these areas the brain may gain instant access to particular parameters regarding the other parameters as fixed and their variation as noise.

Our analysis starts with the measure of dependence between the random variables' retinal speed and direction for a set of positions in the field of view. Then we analyze the properties such as mean, standard deviation, skewness, kurtosis, and negentropy of the distributions of speed and direction. The results provide the most important properties of the distributions depending on the position in field of view. Finally, we investigate the statistical dependencies between the distributions of optic flow and the distributions of depth in the scene, depth of the fixation point, and heading. To put the influence of heading and scene structure into context, the same analysis is performed with two other sets of optic flow fields. One set is generated under the assumption that no gaze stabilization is executed and therefore no rotation occurs. The second set provides a baseline for comparison of the influence of the scene structure. In this set, the depth values are randomly mixed. Thus, the scene structure is abolished and the depth statistics do not differ for different positions.

Methods

Construction of retinal flow fields

In this section, we describe the preparation of retinal flow fields in a sufficient number for the statistical analysis. The calculation of retinal optic flow fields relies on the knowledge of the depth map of a variety of natural scenes. We will explain how to obtain ego-motion parameters and how to construct flow fields from the depth map and the ego-motion.

We generate flow fields under three different conditions. One condition is regarded as naturalistic and combines naturalistic ego-motion, which includes gaze stabilization, through natural scenes (natural condition). Another set of flow fields relies on the same set of natural scenes and heading directions but without gaze stabilization (nonstabilized condition). In this set, gaze is directed to the same objects in the visual field as in the natural condition but is not stabilized on that object, i.e., does not conteract the motion induced by the forward movement. The third set of flow fields is generated from the same naturalistic ego-motion parameters, including gaze stabilization, but each scene is mixed in depth by exchanging the depth values between randomly selected pairs of positions (mixed depth condition). This procedure ensures that the overall distribution of depth values is natural, but the differences in depth statistics for different positions in the visual field disappears.

Database. We use the Brown range image database, a database of 197 range images collected by Ann Lee, Jinggang Huang and David Mumford at Brown University (Huang et al. 2000). The range images are recorded with a laser range-finder with high spatial resolution. Each image contains 444×1440 measurements with an angular separation of 0.18°. The field of view is 80° vertically and 259° horizontally. The distance of each point is calculated from the time of flight of the laser beam, where the operational range of the sensor is 2–200 m. The wavelength of the laser beam is 0.9 μm and lies in the near infrared region. Thus, the data of each point consist of 4 values: the distance R, the azimuth angle ϕ, the zenith angle θ, and a value for the reflected intensity of the laser beam. The location of the source of the laser beam is 1.5 m above the ground. Figure 1 shows a typical range-image projected onto the $\phi - \theta$ plane. It can be seen that the intensity of the reflected laser beam characterizes the properties of the reflecting surfaces sufficiently well. The objects in the scene are clearly visible, and the image resembles a grey level picture of a fully illuminated scene at night. The data are provided in spherical coordinates R, ϕ, θ. The 3-dimensional Euclidian coordinates from the standpoint of the laser range finder can be easily calculated by $(X, Y, Z) = (R\cos(\phi)\sin(\theta), R\cos(\theta), R\sin(\phi)\sin(\theta))$.

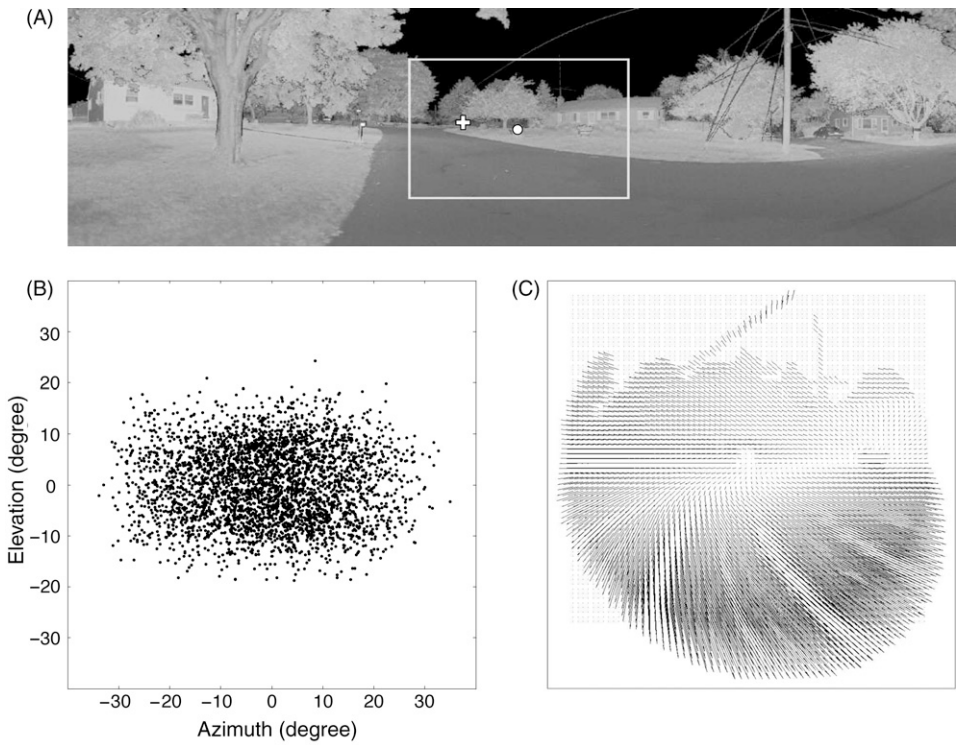

Figure 1. (A) Panoramic projection of 3D data of a range-image. The grey values encode the intensity of the reflected laser beam. (B) Measured gaze directions projected onto the azimuth–elevation plane. (C) Retinal flow field generated by a leftward motion and a gaze-stabilizing eye movement through the scene depicted with the white frame in A. The motion direction is depicted by a cross. The direction of gaze is depicted by a disc.

Retinal projection. The knowledge of the 3D coordinates of each image point allows the calculation of the true motion of that point for any given combination of translation and rotation of the projection surface. As we are interested in the statistics of retinal projections, we consider as the retina a spherical projection surface with the radius 1. All coordinate systems we will use in the following are attached to the center of the projection surface and therefore the coordinates of the data delivered by the database have to be transformed in the perspective of the projection surface. In Euclidian coordinates, the X- and Y-axis are right and up, and the Z-axis is perpendicular to the X–Y plane. The value of the Z- coordinate of any point in the scene is the depth of that point from the perspective of the projection surface. The most simple description of optic flow vectors on the sphere is given by the following notation. Let ϵ be the angle of eccentricity describing the meridians of the sphere and σ the rotation angle describing the circles of latitude rotating counter clockwise. The focal point is defined by $\epsilon = 0$. The meridians and the circles of latitude are coordinate lines, and every vector v on the sphere has the components $v = (v_\epsilon, v_\sigma)$ in the respective local orthonormal coordinate system. The velocity v of a point moving over the sphere described in terms of the temporal derivatives of ϵ and σ is $v = (\mathrm{d}\epsilon/\mathrm{d}t, \sin(\epsilon)\,\mathrm{d}\sigma/\mathrm{d}t)$. Although the spherical coordinates ϵ, σ already sufficiently provide the description of the sphere, we want to use a second spherical coordinate system to denote positions on the projection surface, in terms of which we are going to plot our results. Each position on the sphere is described by the azimuth $\tilde{\phi}$ and the elevation $\tilde{\theta}$, where the projection of a point in the scene onto the sphere is governed by the relationship in equation

$$\frac{X}{Z} = \frac{\sin(\tilde{\phi})}{\cos(\tilde{\phi})} = \frac{\sin(\epsilon)}{\cos(\epsilon)}\cos(\sigma)$$
$$\frac{Y}{Z} = \frac{\sin(\tilde{\theta})}{\cos(\tilde{\phi})\cos(\tilde{\theta})} = \frac{\sin(\epsilon)}{\cos(\epsilon)}\sin(\sigma).$$

Since, positions of the upper and the right visual field are denoted by positive values of $\tilde{\theta}$ and $\tilde{\phi}$ respectively and positions of the lower and left visual field are denoted by negative values of $\tilde{\theta}$ and $\tilde{\phi}$ respectively, the reader can easily discern what positions on the sphere are pointing to the right, left, up, and down from the perspective of the observer. The flow field emerging on a moving sphere can be easily extracted by a simple transformation from the well–known flow field (v_x, v_y), which would be generated on a moving plane with internal coordinates $(x, y) = (X/Z, Y/Z)$ (Longuet-Higgins and Prazdny 1980). The flow field generated on a plane is described by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v_x = \frac{1}{Z}(-T_x + xT_z) + \left(xy\Omega_x - (1+x^2)\Omega_y + y\Omega_z\right) \tag{1}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = v_y = \frac{1}{Z}\left(-T_y + yT_z\right) + \left(-xy\Omega_y + (1+y^2)\Omega_x - x\Omega_z\right). \tag{2}$$

The transformation rule is

$$v_\epsilon = \cos^2(\epsilon)(\cos(\sigma)v_x + \sin(\sigma)v_y) \tag{3}$$

$$v_\sigma = \cos(\epsilon)(\cos(\sigma)v_y - \sin(\sigma)v_x). \tag{4}$$

Ego-motion parameters. To calculate the flow field from the scene structure we need the motion parameters of the projection surface. The ego-motion of the surface is fully described by the translational velocity vector of the surface $T = (T_x, T_y, T_z)$ and the vector of rotation $\Omega = (\Omega_x, \Omega_y, \Omega_z)$ in the Euclidian coordinate system attached to the projection surface. The translation T of the surface can be further split up in the parameters translational or walking speed $\|T\|$ and heading (H_ϕ, H_θ), which are azimuth and elevation denoting the direction of the translational velocity vector of the surface:

$$T = (T_x, T_y, T_z) = \|T\|(\cos(H_\phi)\sin(H_\theta),\ \cos(H_\theta),\ \sin(H_\phi)\sin(H_\theta)).$$

Natural ego-motion within the scenes involves eye movements which stabilize the gaze on environmental objects (Lappe et al. 1998; Lappe 2000;). Gaze stabilization keeps the point of interest or the gaze-attracting object in the center of the visual field and causes the motion in the center of view to be zero. It can be easily extracted from Equations 1 and 2 that the associated rotation depends on the translation by $\Omega = (1/Z_f)(Ty, -T_x, 0)$, where Z_f denotes the depth of the point at which gaze is directed. Under this assumption, (3) and (4) can be transformed to

$$v_\epsilon = \frac{1}{Z}\left(\cos(\epsilon)\sin(\epsilon)T_z + \left(\frac{Z}{Z_f} - \cos^2(\epsilon)\right)\left(\cos(\sigma)T_x + \sin(\sigma)T_y\right)\right) \tag{5}$$

$$v_\sigma = \frac{1}{Z}\cos(\epsilon)\left(\frac{Z}{Z_f} - 1\right)(\cos(\sigma)T_y - \sin(\sigma)T_x). \tag{6}$$

By (5) and (6) the parameters governing the optic flow at a certain position are the walking speed $\|T\|$, the heading (H_ϕ, H_θ) and the depth structure of the scene determined by the depth Z of the point in question and the depth of the fixation point Z_f.

For the condition without gaze stabilization the concerning retinal flow can be extracted from (5) and (6) by assuming the observer gazes towards a point in infinity, i.e. $Z_f \to \infty$, and thus Ω and the term Z/Z_f vanish.

Since, for higher walking speed, the distribution is linearly shifted to higher speed values, there are only trivial correlations between the motion signals and walking speed. Therefore, we restrict our analysis to flow fields generated by a walking speed of $\|T\| = 1.4$ m/s.

Finding plausible ego-motion parameters (H_ϕ, H_θ) and depth of fixations Z_f for the respective scene requires to search for feasible walking directions within the scene and to extract probable gaze directions. The walking direction within the scene combined with the gaze direction provides the parameters of heading (H_ϕ, H_θ). Furthermore, if the direction of fixation is given, the depth of fixation can be extracted from the point in the laser range image that the gaze direction is associated to.

To determine possible walking directions within a range image, we search for areas which are free from obstacles in a depth of at least 3 m and a width of 0.7 m. This criterion gives us a set of walking directions for each scene, which are considered to be equally likely.

To obtain gaze directions that we can use to generate gaze stabilization movements, we measured eye movements of observers who viewed images, that

were generated from segments of the range images centered on the walking directions. The images are projected onto a 36.5×27.5 cm plane with a focal length of 30 cm (white frame in Figure 1A). Six subjects viewed these pictures on a 17 inch computer monitor with the head stabilized on a chin rest 30 cm in front of the monitor. Pictures were shown for 1 s in immediate succession to give the impression of a changing environment the subject is moving through. Gaze fixation points were measured by an eye tracking system (Eye Link II). The first fixation for each picture was rejected because it might be partially driven from the preceding picture. The subsequent fixations were used as probable gaze directions for the statistical analysis. Although, the subjects are not actually walking through the real scenes, and therefore have no access to the true color, disparity and other factors, which might influence gaze attraction, the arrangement of objects and surfaces populating the scene and the objects itself are well recognizable (Figure 1A). Furthermore, humans are usually familiar with that sort of scenes and this world knowledge ensures that the scenes are instantly identifiable as street scenes or forest scenes, and that gaze is instantly attracted to the usual objects in such scenes. Figure 1B shows the distribution of gaze directions while viewing the scenes. The points are plotted in spherical coordinates (ϕ, θ) and are centered on the direction of walking used for the flow field calculations.

To consider all aspects of human walking we also take bouncing and swaying of the head during walking into account. Bouncing and swaying of the head while walking is part of the complex oscillatory motion pattern of during walking (Imai et al. 2001). The position of the head during walking can be modeled as sinusoidal time functions. To obtain typical values for the amplitude and period we measure the head position of one human subject while walking. The walking velocity of the subject was 1.4 m/s. The height of the subject was 1.80 m. The head position was measured by a position tracking system (Motion Star). We approximate the properties of vertical and horizontal movement of the head as follows. The vertical head position has a period of 0.6 s and an amplitude of 0.02 m. The horizontal head position has a period of 1.2 s and an amplitude of 0.02 m as well. The velocity of the head during walking can be obtained by the first temporal derivative of the horizontal and vertical head position. The actual horizontal and vertical movement of the head for a certain motion situation is picked up at a randomly selected time and is added to the preliminary determined ego-motion parameters such that the gaze of the eye towards the fixation point is stabilized also during bouncing and swaying of the head.

Since the actual combination of walking direction, gaze direction and head movement is linked to the environment and to the task, and since the simulated components of ego-motion might not be generally independent, our assumed ego-motion is an approximation to actual ego-motion and might not match actual ego-motion in all details. But the simulation matches the main components of human ego-motion and allows us to combine naturalistic ego-motion parameters with the true depth information data provided by the range image database.

We mirror each scene and the respective heading on the vertical plane (Y–Z plane). This procedure of mirroring the scene attaches to each position in the field of view the depth value of its counterpart in the opposite hemisphere and therefore doubles the set of depth data points for each position.

Simultaneously mirroring the heading ensures that ego motion is still in the direction of the obstacle-free corridor.

The retina. The flow fields we finally consider are elicited on the inside of a section of a sphere, in which a grid of motion sensors is affixed. The field of view of this retina is set to 90° horizontally and 58° vertically. This field of view is subdivided in pixels, which can be referred to as motion sensors, with a resolution of $0.36^\circ \times 0.36^\circ$ yielding a grid of 250×160 pixels. As the angular separation of the range images is 0.18°, one pixel covers up to four data points. The depth values $Z = R \sin(\phi) \sin(\theta)$ from these data points are averaged. The mean depth value is assigned to the pixel in question. Thus, we reduce the original resolution provided by the laser range images. This procedure is necessary, because the pixel grid of the retina is sliding over the pixel grid of the laser range image and mostly does not match the original pixels. Therefore, reducing the resolution ensures that all pixels of the retina receive appropriate motion signals. The flow vector attached to a pixel depends on the depth value, the translation and rotation components of the ego-motion and the visual field position of the pixel. The flow vectors of all pixels of an image provide the measurement of the true retinal flow field for this gaze direction, ego-motion, and scene. Figure 1C shows an example of a true retinal flow field.

Statistical analysis

We constructed 7136 different flow fields for each of the three conditions: naturalistic, no gaze stabilization, and mixed depth map. To examine the local statistics of these flow fields we collect for each scene, motion situation, and pixel position, the data sets which comprises the retinal velocity, the depth at the respective position, the depth of fixation, and the heading of the respective ego-motion. Examples of the distributions of retinal velocities can be seen in Figure 2.

The analysis of the local statistics of optic flow is divided into two parts. The first part is dedicated to the examination of the distributions of optic flow and how strong the statistical properties of which depend on the positions in the field of view. The investigation considers the polar optic flow components' retinal flow direction and retinal speed and starts with the measure of the statistical dependence of these two variables. Although there are slight differences in the degree of statistical dependence for the different conditions, it turns out that flow direction and retinal speed are largely statistically independent for all conditions. Therefore, the further analysis is based on the extracted 1-dimensional distributions of retinal speed and retinal flow direction. For each position in the field of view we measure the mean, the scatter, the skewness, the kurtosis, and an estimation of the negentropy for the distributions of retinal speed and retinal flow direction.

The second part addresses the problem of the statistical dependence of the variables' retinal direction and retinal speed on the particular parameters' depth, depth of fixation point, and heading. The statistical correlations are not purely linear, and nonlinear statistical dependencies play an important role. This can be seen in the exemplary scatter plots (Figure 3), which show the extracted data for the

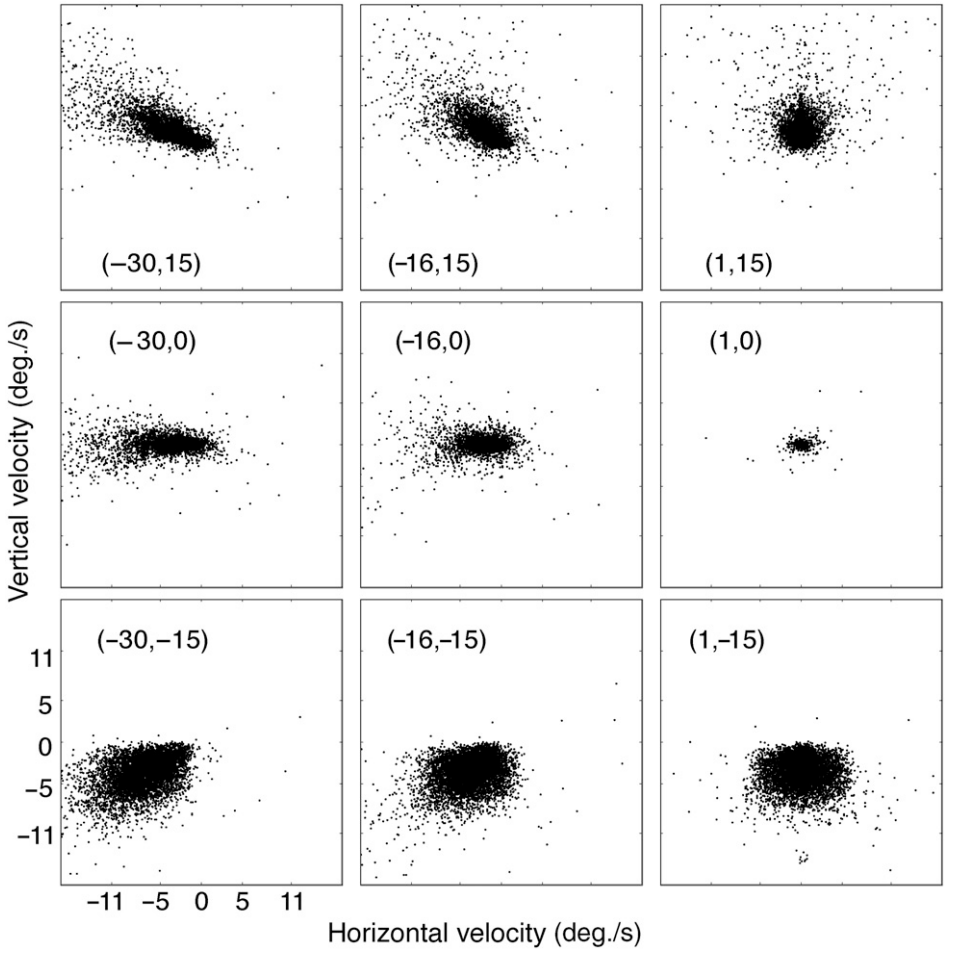

Figure 2. Measured distributions of retinal velocity for 9 different positions in the left visual field. The numbers in each panel give the visual field position in spherical coordinates $(\tilde{\phi}, \tilde{\theta})$ in degrees. First row: positions in the upper field of view, middle row: horizontal meridian, third row: lower field of view.

distribution of retinal speed and the inverse of depth (Figure 3A), the retinal speed and the elevation heading component (Figure 3B), and the retinal speed and the azimuth heading component (Figure 3C) at the visual field position $(-5^\circ, -18^\circ)$. All scatter plots show that a statistical analysis based on linear correlation is not sufficient. For example in the scatter plots Figure 3A and C, a quadratic correlation seems to underlie the data set and in the scatter plot Figure 3B a correlation of degree 3 provides the main contribution. However, despite the very different kinds of statistical dependencies for different data sets and for different positions we would like to compare the degree of dependence the retinal flow has on the different parameters. Information theory provides the notion of mutual information between random variables. The mutual information is a measure of the difference between the joint probability density function (PDF) of the random variables and the PDF

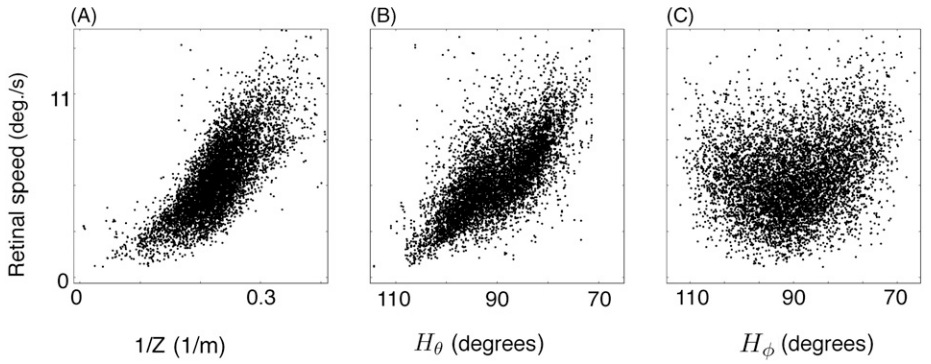

Figure 3. Examples of scatter plots of the dependence of retinal speed on inverse depth (A), elevation of heading (B), azimuth of heading (C). The data are from the visual field position $(-5°, -18°)$.

which would appear if the random variables were statistically independent. Since the mutual information only vanishes if the random variables are fully statistically independent, the measure of the mutual information takes all kinds of statistical dependencies into account and is useful to investigate the degree of dependence between random variables. However, the mutual information can range from zero to infinity. Therefore, we are going to define a generalized dependence coefficient in terms of the mutual information between data sets which will be bound between values 0 and 1. This definition is motivated by the mathematical relation between mutual information and the linear correlation in the case of purely linear statistical dependency.

Polar representation of retinal flow vectors. The retinal velocity at a certain position in the field of view is a 2-dimensional vector (v_ϵ, v_σ). Therefore, the velocity distributions are distributions of 2-dimensional random variables. The random variables we choose for further analysis are the polar coordinates retinal speed v and retinal direction ϕ_{dir}. We will show that these random variables are largely independent. Thus, our analysis of the local statistics of retinal velocities will be separately performed on these two random variables:

$$v = \sqrt{v_\epsilon^2 + v_\sigma^2},$$

$$\phi_{\text{dir}} = \begin{cases} \arccos \dfrac{v_\epsilon}{\sqrt{v_\epsilon^2 + v_\sigma^2}} & ; \quad v_\sigma \geq 0, \\ -\arccos \dfrac{v_\epsilon}{\sqrt{v_\epsilon^2 + v_\sigma^2}} & ; \quad v_\sigma < 0 \end{cases}.$$

Estimating the properties of the distributions of speed and direction. To illustrate the different properties of the distributions of speed and direction for different positions

in the visual field we measure the mean (EX), the scatter (DX), the skewness (M_3X), the kurtosis (M_4X), and estimate the negentropy ($\mathcal{J}X$). For the sake of completeness, we list the well-known corresponding formulas to estimate these parameters from an empirical data set $X = \{x_i \in \mathbb{R}\}_{i=1,2,\ldots,N}$:

$$EX = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$DX = \overline{M_2X} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - EX)^2}$$

$$M_3X = \frac{(1/N-1) \sum_{i=1}^{N} (x_i - EX)^3}{DX^3}$$

$$M_4X = \frac{(1/N-1) \sum_{i=1}^{N} (x_i - EX)^4}{DX^4} - 3.$$

DX is a measure for the width or the spreading of the data, M_3X and M_4X are measures of the difference between the empirical distribution and the Gaussian distribution. M_3X measures the asymmetry and M_4X is a measure for the flatness of the distribution. A negative value of M_4X means the distribution is more flat than a Gaussian and has shorter tails. A positive value of M_4X means the distribution has a peak higher than a Gaussian and longer tails.

Mean, scatter, skewness and kurtosis take only the first four moments of a distribution into account. To obtain a compact measure to assess the difference between an empirical distribution and the Gaussian distribution with the same mean and scatter we estimate the negentropy from the empirical distribution. Let $P(x)$ be the PDF of a random variable X. The negentropy $\mathcal{J}(X)$ is defined as the difference between the differential entropy of the Gaussian $H(X_{\text{gauss}})$ and the actual differential entropy $H(X) := -\int_{\text{supp}P} P(x) \log_2(P(x)) \mathrm{d}x$

$$\mathcal{J}(X) = H(X_{\text{gauss}}) - H(X) = \frac{1}{2} \log_2(e 2\pi DX^2) - H(X),$$

where e is the Euler number. Since a Gaussian distribution has the maximal entropy for a given mean and scatter, $\mathcal{J}(X)$ is always nonnegative and vanishes only if X is a Gaussian distribution. The estimation of the negentropy of a distribution requires the estimation of the differential entropy.

The measure of dependence and the estimation of differential entropy. The analysis of the statistical interdependence between the optic flow components' retinal direction and retinal speed and the statistical dependencies of the optic flow components on the parameters' depth, depth of fixation point and heading components requires estimating the mutual information from 2-dimensional and 3-dimensional data sets. The analysis of the properties of the distributions of retinal speed and retinal direction furthermore needs an estimation of the differential entropy from 1-dimensional data sets.

The mutual information $mI(X_1, X_2, \ldots, X_n)$ between (possibly more-dimensional) continuous random variables X_i with minor PDFs $P_i(x_i)$ and joint

PDF $P_X(x_1, x_2, \ldots)$ is defined by

$$mI(X_1, X_2, \ldots) = \int_{\mathbb{R}^m} P_X(x_1, x_2, \ldots) \log_2\left(\frac{P_X(x_1, x_2, \ldots)}{P_1(x_1)P_2(x_2)\ldots P_n(x_n)}\right) \mathrm{d}^m x,$$
$$= \sum_i^n H(X_i) - H(X_1, X_2, \ldots), \tag{7}$$

where $m = \sum_{i=1}^n \dim(X_i)$ and

$$H(X) = -\int P_X(x) \log_2(P_X(x) \mathrm{d}^m x; \quad m = \dim(X) \tag{8}$$

is the (differential) entropy for the random variable X. The mutual information is always nonnegative, and zero if and only if $P(x_1, x_2, \ldots) = P_1(x_1)P_2(x_2)\ldots P_n(x_n)$, i.e. the X_i are mutually independent random variables. In this study, we deploy the method of the k-nearest neighbors distances to estimate the differential entropy (Kozachenko and Leonenko 1987) and the modification to estimate mutual information (Kraskov et al. 2004). Let $\Psi = \{\psi_i\}_{i=1, 2, \ldots, N}$ be an m-dimensional data set and let $\hat{P}_\Psi$ be an estimator for the actual PDF P_Ψ. The differential entropy (8) can be estimated by

$$H(\Psi) \approx \hat{H}(\Psi) = -\frac{1}{N} \sum_{i=1}^N \log_2(\hat{P}_\Psi(\psi_i)). \tag{9}$$

Let ϵ_i be the minimal radius of the sphere centered at $\psi_i \in \Psi$ and located within the k nearest neighbors of ψ_i; for large N and large k (but $k \ll N$, $P_\Psi P(\psi_i)$ can be estimated by

$$\hat{P}_\Psi(\psi_i) = \frac{1}{V_u \epsilon_i^m} \frac{k}{N}, \tag{10}$$

where V_u is the volume of the unit ball. Equation 10 leads directly to an estimate of the differential entropy

$$\hat{H}(\Psi) = \log_2(N) - \log_2(k) + \log_2(V_u) + \frac{m}{N} \sum_{i=1}^N \log_2(\epsilon_i). \tag{11}$$

For smaller N and/or smaller k Equation 10 gives a rather bad estimate of P_Ψ and Equation 11 must be replaced by

$$\hat{H}(\Psi) = (\psi(N) - \psi(k))/\log(2) + \log_2(V_u) + \frac{m}{N} \sum_{i=1}^N \log_2(\epsilon_i),$$

where ψ is the digamma function (Kraskov et al. 2004). Let now $\Psi = (\Psi_1, \Psi_2) = \{(\psi_{1_i}, \psi_{2_i}\}_{i=1, 2, \ldots, N}$ be a $(m_1 + m_2)$-dimensional data set for which we wish to estimate the mutual information $mI(\Psi_1, \Psi_2) \approx \hat{H}(\Psi_1) + \hat{H}(\Psi_2) - \hat{H}(\Psi)$. Whereas $\hat{H}(\Psi)$ can be estimated by Equation 11, to use the same distance scales in the joint and the minor spaces, and to avoid any biases the estimation of the differential entropies $\hat{H}(\Psi_1)$ and $\hat{H}(\Psi_2)$ have to be modified in the following way (Kraskov et al. 2004). Let ϵ_i be the minimal radius of the sphere which is centered at $\psi_i \in \Psi$ and which are located, within the k nearest neighbors of ψ_i then k_{1_i} is the

number of data located within the sphere with radius ϵ_i centered at ψ_{1_i}, in the space Ψ_1, and k_{2_i} is the analog number of data around ψ_{2_i}. The estimation of the differential entropy $\hat{H}(\Psi_1)$ is modified by

$$\hat{H}(\Psi_1) = \log_2(N) + \frac{-1}{N}\sum_{i=1}^{N} \log_2(k_{1_i} + 1) + \log_2(V_{1_u}) + \frac{m_1}{N}\sum_{i=1}^{N} \log_2(\epsilon_i).$$

Analogous calculations are performed for the estimation of $\hat{H}(\Psi_2)$. For the estimation of mutual information between more than two random variables, the method can be easily extended.

For all mutual information estimation performed in this study we fixed the number of nearest neighbors by $k = 0.005N$. Whereas N ranges between 6000 and 7136, the number of nearest neighbors takes values between 30 and 35.

Rank ordering of data sets. The dependent components of an empirical data set are usually measured in different units and have different scales. Large differences in scale can cause errors in the estimation of mutual information. Note that the mutual information (7) is preserved under any differentiable transformation $f : \mathbb{R}^m \to \mathbb{R}^m$ of the m-dimensional components. To conform the scales of the components of the data sets, the components are transformed to a uniform distribution by rank ordering.

Let $P_X(x)$, $x \in \Omega \subseteq \mathbb{R}$ be the PDF of a 1-dimensional random variable X. Let H be the Heaviside step function. The transformation which turns X into a uniform random variable is

$$f_X(x) := \int_{\Omega} H(x - \tilde{x})P(\tilde{x})\mathrm{d}\tilde{x}. \tag{12}$$

Let $\{\psi_i \in \mathbb{R}\}_{i=1,2,\ldots,N}$ be a 1-dimensional data set. Then (12) leads directly to the approximation of the uniforming procedure by

$$f_\Psi(\psi_i) := \frac{1}{N}\sum_{j=1}^{N} H(\psi_i - \psi_j), \tag{13}$$

which is referred to as rank ordering. To perform the uniforming procedure for a 2-dimensional data set the first component is rank-ordered according to Equation 13. The resulting data set can then be divided into stripes of the same width comprising the same number of data. Regarding each stripe as a 1-dimensional data set, the stripes are rank-ordered by Equation 13 again. Although uniforming dissolves the internal dependence structure of the 1-dimensional random variable, the mutual information between the uniformed 2-dimensional random variable and a separately rank-ordered third random variable is not affected.

The definition of the generalized dependence coefficient. Suppose the random variables X and Y have the following joint PDF:

$$P(x,y) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{x^2}{2\sigma_1^2}\right)\exp\left(-\frac{(x-y)^2}{2\sigma_2^2}\right), \tag{14}$$

and P_X and P_Y are the PDFs of the constituents. There is a simple relationship between the linear correlation coefficient $C(X, Y)$ and the mutual information $mI(X, Y) = \int_{\text{supp}P} P(x, y) \log_2(P(x, y)/P_X(x)P_Y(y))\, \mathrm{d}x\, \mathrm{d}y$:

$$C(X, Y)^2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \tag{15}$$

$$mI(X, Y) = \frac{1}{2} \log_2 \left(\frac{\sigma_1^2 + \sigma_2^2}{\sigma_2^2} \right) \tag{16}$$

$$C(X, Y)^2 = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} = 1 - 2^{(-2mI(X, Y))}. \tag{17}$$

Motivated by Equation 17 we define for a multi-dimensional random variable Ψ a normed mutual information

$$mI^{\text{normed}}(\Psi) := 1 - 2^{(-2mI(\Psi))}, \tag{18}$$

which takes values between 0 and 1, and which is referred to as a generalized dependence coefficient. Recall that Equation 17 is only valid for linear correlations and if all constituent random variables are gaussian-distributed. However, we will use the definition (18) to condense the estimated mutual information in a value between 0 and 1 for a more compact presentation.

Results

Statistical interdependencies between retinal direction and retinal speed

Figure 4 shows the estimated normed mutual informations $mI^{\text{normed}}(\Phi_{\text{dir}}, V)$ between the distributions of direction and speed for all positions in the field of view in each of the three conditions. The values of $mI^{\text{normed}}(\Phi_{\text{dir}}, V)$ for the natural condition (Figure 4A) range from 0.04 to 0.19, with the peak in the center of the visual field. These values are rather low suggesting that direction and speed are

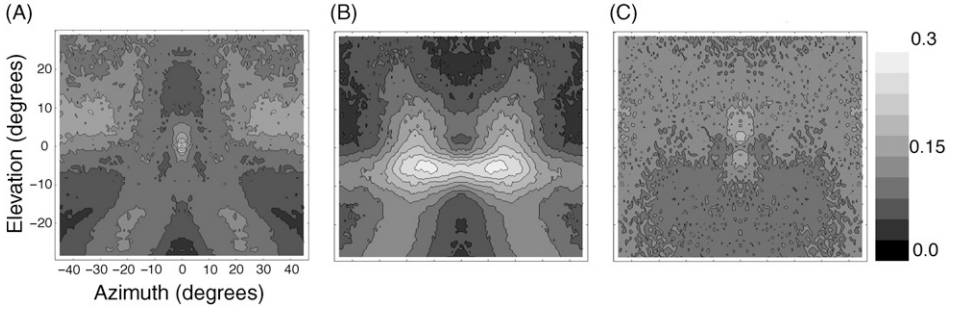

Figure 4. Statistical dependence between retinal speed and direction. Three-dimensional plots and contour plots of the estimated normed mutual informations mI^{normed} (Φ_{dir}, V) between retinal flow direction and retinal speed as function of the position of the visual field for different conditions, (A) natural, (B) nonstabilized and (C) mixed depth.

largely independent from each other at all positions in the visual field. This result, however, is not a direct consequence of Equations 3–6 but rather depends on the statistical properties of the motion and depth parameters and their combination in walking and gaze stabilization. This can be seen from the comparison with the other two conditions.

In the condition with no gaze stabilization (Figure 4B), $mI^{\mathrm{normed}}(\Phi_{\mathrm{dir}}, V)$ ranges from 0.05 to 0.3 in the lower visual field and from 0.04 to 0.2 in the upper visual field. The interdependence between retinal speed and direction is increased for a domain of the lower visual field right under the horizontal line. This shows that the ego-motion situation influences the dependence structure of retinal speed and direction.

In the third condition (mixed depth map, Figure 4C), $mI^{\mathrm{normed}}(\Phi_{\mathrm{dir}}, V)$ varies between 0.08 and 0.16 across the visual field. Thus, randomization of the depth structure keeps the statistical interdependence between retinal speed and direction on the same level as in the natural condition. Thus, the depth structure exerts less influence on the interdependence between retinal speed and direction than the ego-motion situation. However, a depth structure with very different statistics may affect the interdependence between direction and speed. For instance, if the scene contains only objects in great distances from the observer the retinal motion signals are mostly caused by the rotational component of ego-motion. This produces a higher statistical interdependence between direction and speed. Therefore, the low statistical interdependence between direction and speed of the optic flow in the natural condition is a particular property of ego-motion through natural settings.

The statistical interdependence between retinal direction and speed is not dependent on walking speed. Variation of walking speed only scales the retinal speed by a proportionality factor. However, a statistical variation of walking speed between different motion situations would evidently further diminish the level of statistical interdependence between retinal direction and speed by introducing an additional statistical variance which solely affects the statistics of retinal speed.

In the remainder of our analysis, we consider the statistical properties of the distributions of retinal direction and speed separately. This is justified by the low statistical interdependence between direction and speed in the natural condition and facilitates the understanding and interpretation of the results.

Properties of the distributions of speed and directions

Figure 5 shows some examples for kernel-based estimates of the PDFs of direction (measured as deviations from the radial direction) and speed (Parzen 1962; Silverman 1986). All examples are from the natural condition. They show different positions in the left visual field. The distributions of the right visual field are essentially mirror symmetric. The estimated distributions of speed appear similar to logarithmic Gaussian distributions. We therefore measure the skewness, kurtosis, and negentropy for the logarithmic values of speed rather than for the speed itself to reveal how close the speed distributions are to logarithmic Gaussian distributions. We note the reference to the logarithm of a random variable by the prefix log (for example log-kurtosis).

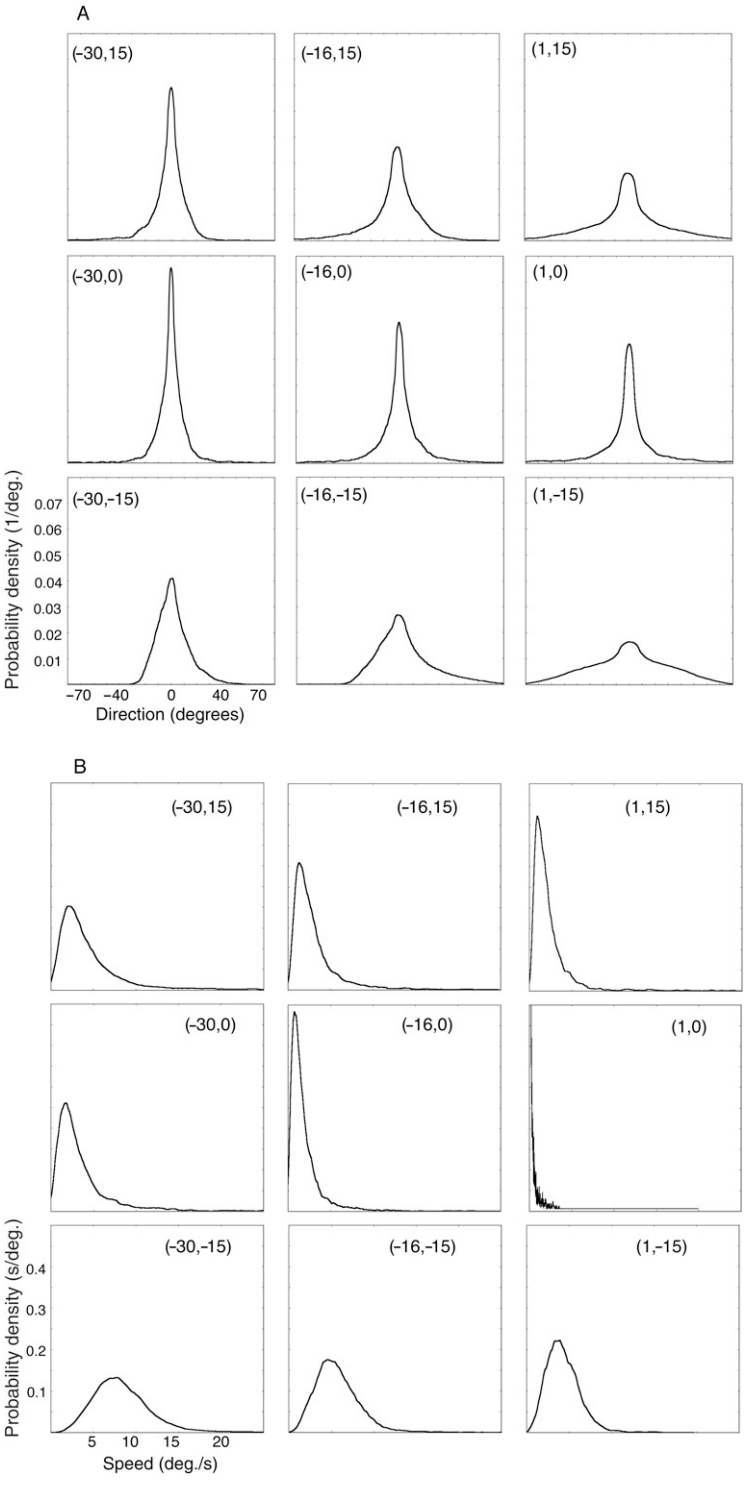

Figure 5. Examples of kernel-based density estimations for direction (A) and retinal speed (B). Nine positions in the left visual field are shown.

Properties of the distributions of directions. Figure 6 shows the visual field maps for mean, scatter, skewness, kurtosis, and negentropy for the distributions of direction in the three conditions. First, we discuss the results concerning the natural condition (column A). The top panel (Figure 6A1) shows the mean of the distributions of direction for all positions in the field of view. The mean deviates

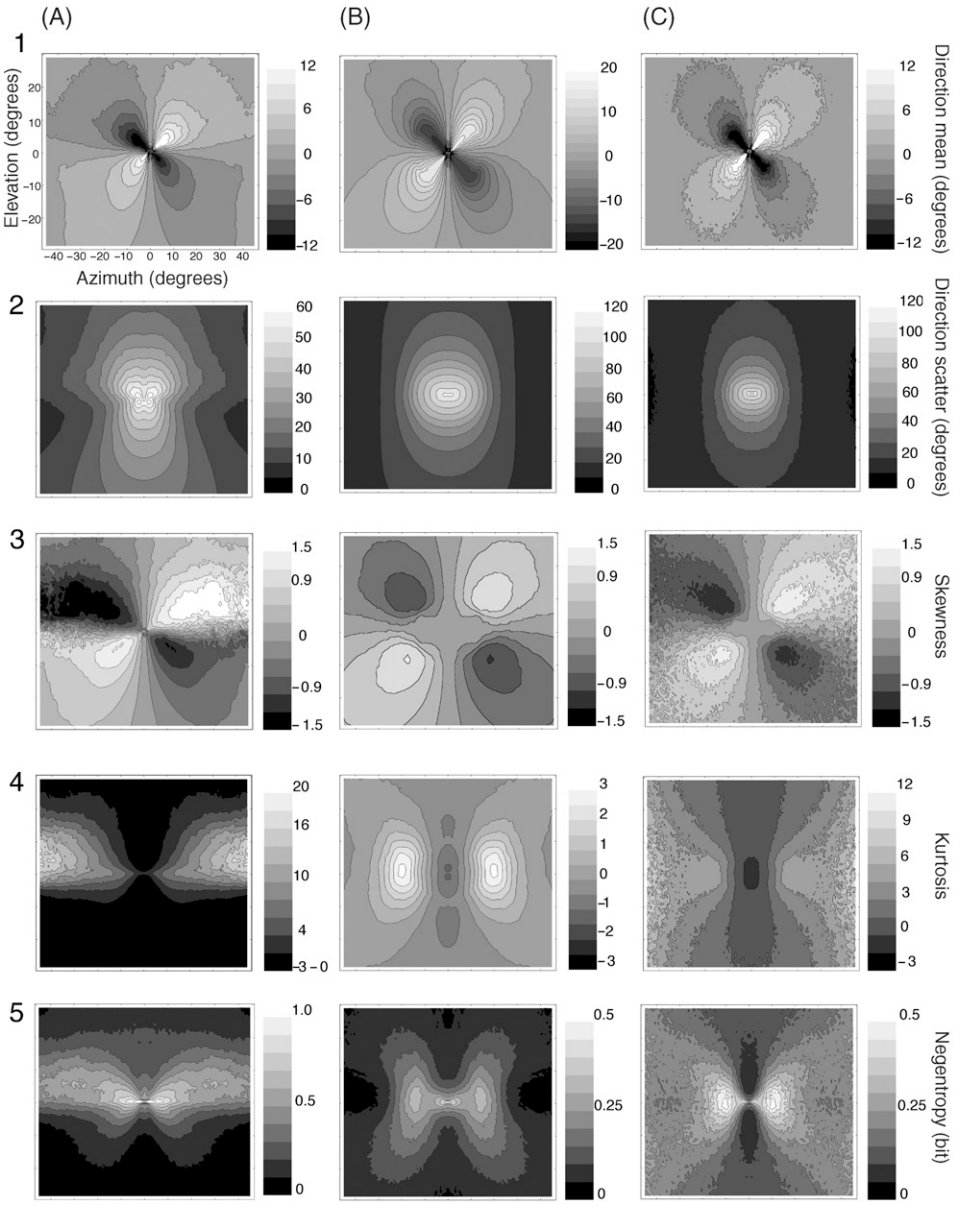

Figure 6. Estimated statistical properties of the distributions of directions for the different conditions plotted over the visual field. 1: Estimated mean, 2: estimated scatter, 3: estimated skewness, 4: estimated kurtosis, 5: estimated negentropy.

from the radial direction by up to 12°. The deviation from radial is high near the center of the field of view and decreases towards the periphery. The variations of the deviation with eccentricity and the absolute values of deviation are very similar in the upper and lower visual fields. The visible cloverleaf structure in the plot shows that the means of the direction in each quadrant are distorted towards the vertical direction: the means are negative in the left upper visual field and positive in the left lower visual field, and vice versa for the right visual field. A similar structure occurs in the map of the skewness of the direction distributions (Figure 6A3). This means that the direction distributions have longer tails at the side where the directions are closer to the vertical direction. Consequently, the skewness vanishes for the positions on the horizontal and vertical meridian.

The shifts of the means and the distortions of the distributions are mainly caused by the interplay between the mathematics of the projection and the properties of the distribution of headings. The distribution of headings has a higher variance for the horizontal component than for the vertical component (cf. Figure 1B). When heading is varied symmetrically around the center along the horizontal meridian the flow vectors induced at position along the 45° diagonal in the lower visual field are distributed asymmetrically around the radial direction (Figure 7). As our flow fields include eye rotations to stabilize gaze on an attended object in the scene, the flow vectors are additionally influenced by the properties of the distributions of Z/Zf (see Equations 5 and 6). Z/Zf takes values between 0 and infinity. The distribution of Z/Zf is right-skewed. The result is a further skewing of the direction distributions. Therefore, the extend to which the resulting distributions of flow direction are skewed depends on the interplay between the statistics of the depth structure and the statistics of the ego-motion parameters. The different depth statistics in the upper and lower visual fields lead to the differences in the magnitude of skewness in the upper and the lower visual field in Figure 6A3.

Figure 6A2 shows the scatter of the directions around the mean. The scatter is maximal (about 60°) at the center of the field of view and decrease to around 10° in

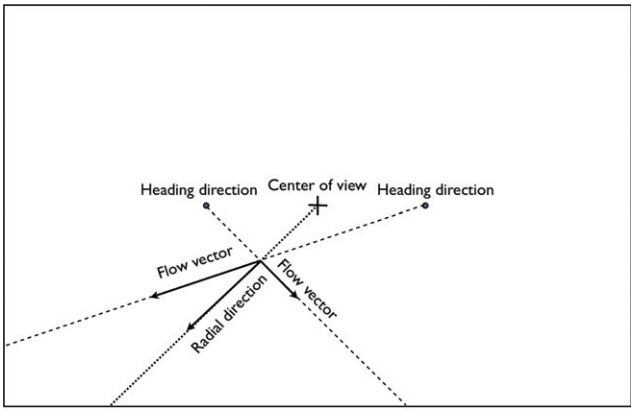

Figure 7. In the case of straight forward ego-motion without rotations, the same horizontal deviation of heading to the left and to the right, respectively, results in different deviations from the radial direction for the resulting flow vectors. The distribution of flow directions is skewed.

the periphery. In combination with the mean the scatter map shows that the direction distributions in the periphery become more radial. There is not much difference between the upper and the lower visual field.

The plots of kurtosis and estimated negentropy (Figure 6A4 and A5) show that in large areas of the lower visual field the kurtosis and negentropy are very small (from −0.4 to 2.0 and 0.02 to 0.08 respectively). Kurtosis and negentropy increase (up to 17 and 1 respectively) near the horizontal meridian. However, also comparatively small deviations from zero kurtosis and from zero negentropy, such as those in the lower visual field, give a significant difference of the distribution from a Gaussian. For example, the distribution at position (−30, −15) in Figure 5A has a kurtosis of 0.44, a negentropy of 0.05, and a skewness of 0.55, and is clearly different from a Gaussian distribution. Position (−30, 15) in Figure 5A provides an example for a distribution with rather large values of kurtosis (6.54), negentropy (0.28), and skewness (−1.28).

We conclude that the distributions of the directions for positions of the lower visual field are rather close to Gaussian distributions and that extreme non-Gaussian distributions occur near the horizontal meridian.

A comparison with the direction distributions in the nonstabilized condition (second column) and the mixed depth condition (third column) shows the influence of the statistics of the ego-motion parameters and the depth on the distributions of retinal directions. Whereas the cloverleaf structure in the mean exists in all conditions, the patterns of scatter, skewness, kurtosis, and negentropy show clear differences between the conditions. In the nonstabilized and mixed depth conditions, there are no differences between upper and lower visual field. Kurtosis and negentropy do not take as high values as some domains of the visual field in the natural condition. Although the distributions in both conditions are similarly skewed, natural motion parameters in combination with natural scenes have a cumulative effect on skewing for some regions of the visual field.

Properties of the distributions of retinal speed. Figure 8 shows the visual field maps for mean, scatter, skewness, kurtosis, and negentropy for the distributions of retinal speed for the three conditions. Figure 8A1 shows the mean of retinal speed for all positions in the field of view in the natural condition. Mean retinal speed and scatter is zero at the center of view because of the assumed gaze stabilization. Mean retinal speed increases in the periphery up to 20°s. Note that the absolute values of speed would scale with walking speed, which was a constant 1.4 m/s in our calculations, but only by a constant factor for all flow speeds. Hence, walking speed does not change the distribution over the field of view.

The increase of the mean speed is larger for the lower visual field than for the upper visual field. The scatter ranging from 0° to 11°/s, on the other hand, increases more in the upper visual field than in lower visual field. These differences between mean and scatter show that the retinal speeds are faster and more uniform in the lower visual field and slower and more variable in the upper visual field.

Since the appearance of the estimated distributions for speed suggests that these distributions are Gaussians on a logarithmic scale, we measured the skewness, kurtosis, and negentropy for the logarithms of speed. This measures show how close the speed distributions are to logarithmic Gaussian distributions. Figure 8, Panels A3, A4 and A5, show the estimated values for the log-scatter, log-kurtosis

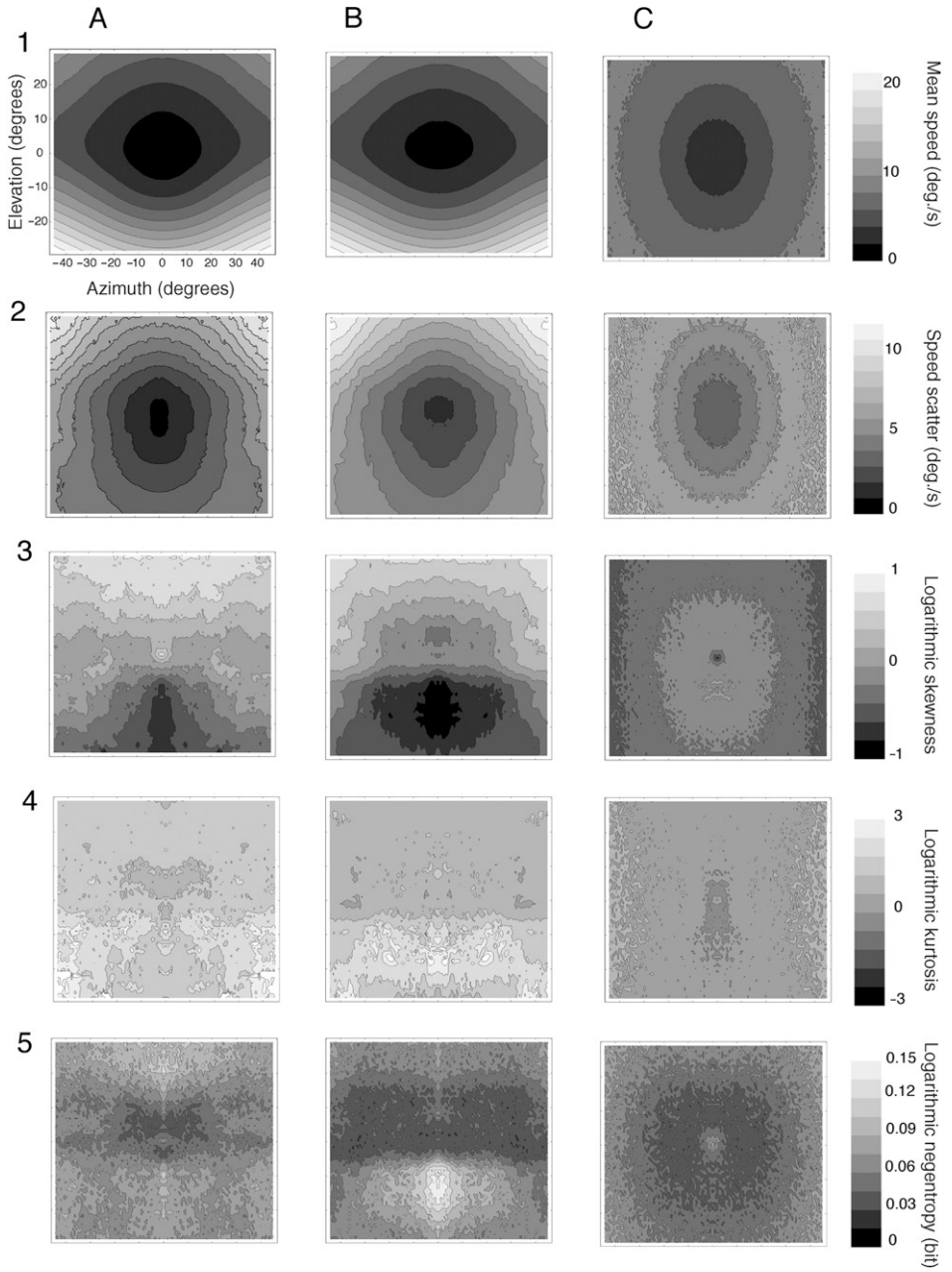

Figure 8. Estimated statistical properties of the distributions of retinal speed for the different conditions (A) natural, (B) nonstabilized and (C) mixed depth, plotted over the visual field. 1: Estimated mean, 2: estimated scatter, 3: estimated skewness of log-speed, 4: estimated kurtosis of log-speed, 5: estimated negentropy of the distributions of log-speed.

and log-negentropy. These values are largely uniform over the visual field. The log-skewness ranges from −0.8 to 0.8, the log-kurtosis from 0.5 to 3, and the estimated log-negentropy from 0.03 and 0.11. Although these values are rather low, for each position either the skewness, or the kurtosis, or both values, are significantly different from zero, which we tested by calculating the standard errors and using the resulting error bars (approx. two times the standard error) as significance criterion. Therefore, the distributions of retinal speed are significantly different from log-Gaussian distributions. However, the small values of log-skew, log-kurtosis, and log-negentropy may, for practical purposes, allow to model the distributions of retinal speed by log-Gaussian distributions.

The distributions in the nonstabilized condition (Figure 8, column B) look very similar to those for the natural condition except for positions close to the center of view. Close to the center of view the mean and scatter of retinal speed do not vanish in the nonstabilized condition and do not fall below 1.3°/s and 1.6°/s, respectively. In the mixed depth condition, the distributions show a complete different pattern (Figure 8, column C). The increases of the mean and scatter towards the periphery are not as pronounced as in the other conditions and do not rise about 8°/s and 7°/s, respectively. Log-skewness takes only negative values across the visual field. Log-kurtosis is smaller than in the other conditions. Log-negentropy is in the same range as in the natural and nonstabilized condition. These results suggest that the depth statistics has a shaping effect on the statistics of retinal speed while the gaze stabilization reflex affects the statistics of retinal speed only in the center of the visual field.

Dependence of the local optic flow statistics on scene structure and ego-motion

In this section, we describe the statistical dependencies of the retinal flow on the depth statistics of the scene (depth Z and fixation depth Z_f) and on the heading direction (H_ϕ, H_θ). The dependence on scene statistics is interesting because the speed of an element of the optic flow depends on the distance of the element from the observer. Moreover, in case of combined observer translation and gaze stabilization the speed and the direction of the motion vector of the optic flow element depends on the relationship between the distance of the element from the observer and the depth of the gaze point. Without gaze stabilization, the statistics of the retinal direction solely depend on the statistics of the heading direction and not on the statistics of depth. Retinal speed is only influenced by depth and not by the depth of the gaze point.

Dependence of the local optic flow statistics on the depth statistics of the scene. To reveal to what extend the optic flow in the natural condition is statistically dependent on the depth statistics of the natural environment, we estimated the normed mutual information between the random variables' retinal flow direction Φ_{dir} and speed V and the following random variables of the scene structure: the 2-dimensional vector $(1/Z, 1/Z_f)$, the inverse depth-values $1/z$, and the quotient between the depth and the depth of the fixation point Z/Z_f. We take Z/Z_f rather than $1/Z_f$ as a single random variable because the direction of retinal motion depends on Z/Z_f not on Z_f alone.

Figure 9 column A shows the distribution of the normed mutual information over the visual field for the aforementioned parameter combinations.
The estimated values for $mI^{\text{normed}}(\Phi_{\text{dir}}, (1/Z, 1/Z_f))$ (Figure 9A1) range from 0.05 to 0.6. Positions in the lower visual field show smaller values than positions in the

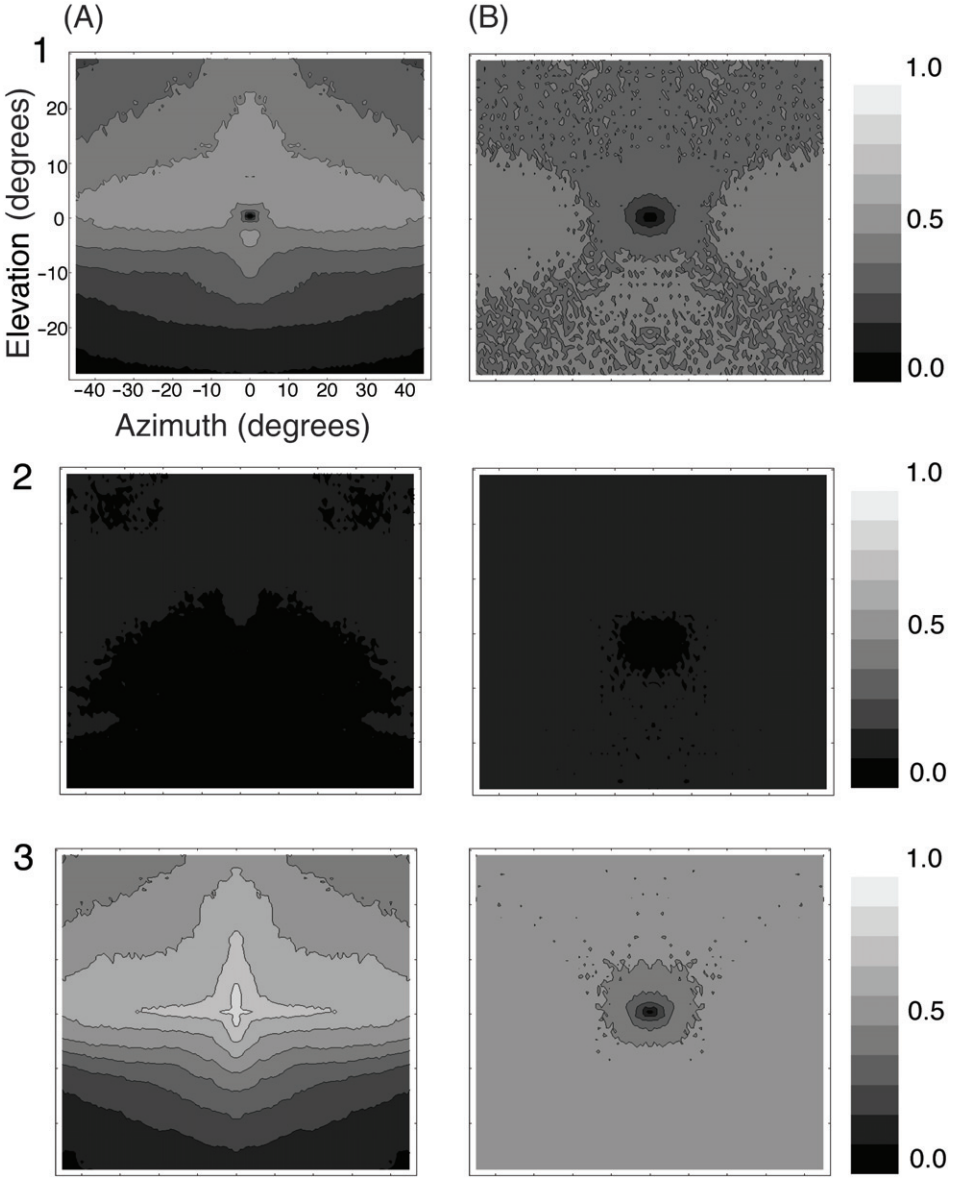

Figure 9. Statistical dependence between direction and depth. Estimated normed mutual information in the natural (A) and the mixed depth (B) condition plotted over the visual field. 1: mI^{normed} $(\Phi_{\text{dir}}, (1/Z, 1/Z_f))$ between retinal direction Φ_{dir} and depth structure $(1/Z, 1/Z_f)$, 2: mI^{normed} $(\Phi_{\text{dir}}, 1/Z)$ between Φ_{dir} and the inverse of depth $1/Z$, 3: mI^{normed} $(\Phi_{\text{dir}}, Z/Z_f)$ between Φ_{dir} and the quotient of depth and fixation depth Z/Z_f.

upper visual field. The highest values are observed along the horizontal meridian. The estimated values for $mI^{\text{normed}}(\Phi_{\text{dir}}, 1/Z)$ (Figure 9A2) are very small and nowhere exceed 0.16. The estimated values for $mI^{\text{normed}}(\Phi_{\text{dir}}, Z/Z_f)$ (Figure 9A3) show a similar distribution as $mI^{\text{normed}}(\Phi_{\text{dir}}, (1/Z, 1/Z_f))$, but with a peak of higher dependence (0.8) in the center of view. Taken together, these plots show that the dependence of the direction of the optic flow on the depth statistics of the scene is particularly strong in the upper visual field and that the combination of depth Z and fixation depth Z_f considerably influences the flow directions in this area. In contrast, the flow directions in the lower visual field do not carry much information about scene structure. Finally, the statistics of $1/Z$ by itself has hardly any influence on the statistics of the retinal direction (Figure 9A2).

In the mixed-depth condition (Figure 9, column B), the dependence of direction on depth shows only minor variation over the visual field. The differences between upper and lower visual field in the natural conditions disappear in the mixed-depth condition. The statistical dependence of depth ($1/Z$) on the statistics of flow directions remains minor. The nonstabilized condition is not shown, because in this condition retinal direction does not depend on the depth structure of the scene, and all values would be zero.

Figure 10 shows the statistical dependence of retinal speed on the depth structure of the scene. Compared to retinal direction (Figure 9A), retinal speed shows an almost opposite dependence on the statistics of depth in the scene in the natural condition (Figure 10, column A). The estimated values for $mI^{\text{normed}}(V, (1/Z, 1/Z_f))$ vary between 0.6 and 0.97 (Figure 10A1), with smaller values in the lower visual field and high values in the upper visual field. The estimated values for $mI^{\text{normed}}(V, 1/Z)$ (Figure 10A2), range from 0.58 to 0.97 and show a visual field distribution similar to that of $mI^{\text{normed}}(V, (1/Z, 1/Z_f))$. The estimated values for $mI^{\text{normed}}(V, Z/Z_f)$ (Figure 10A3) range between 0.05 and 0.5. They show a minor influence of Z/Z_f on the statistics of retinal speed for the lower visual field and a moderately increased statistical dependence for the upper visual field. The latter is caused by an increased statistical dependence between $1/Z$ and Z/Z_f for the upper visual field (data not shown). We therefore conclude that the dependence of the distributions of retinal speed on the depth statistics of the scene is mainly carried by $1/Z$.

In the nonstabilized condition, retinal speed depends only on $1/Z$. The estimated statistical dependence of retinal speed on $1/Z$ resembles the estimated data in the natural condition (Figure 10B2). In the mixed-depth condition, retinal speed is highly dependent on depth at all positions of the visual field but the dependence on depth is symmetric between upper and lower field.

We find that considering scenes with a nonnatural depth statistics results in completely modified statistical dependencies between depth and retinal optic flow. Thus, we conclude that these dependencies are specific in the case of natural scenes.

Dependence of the local optic flow statistics on the statistics of heading. The dependence of the statistics of the retinal direction on the statistics of the observer's heading is shown in Figure 11. Column A depicts the natural condition. The estimated values for $mI^{\text{normed}}(\Phi_{\text{dir}}, (H_\phi, H_\theta))$ (Figure 11A1), range between 0.5 and 0.99 and thus reveal a strong statistical influence of heading on the flow direction. This influence is much more pronounced in the lower visual field than in the upper visual field. This observation is in accordance with the observation in the previous section that the

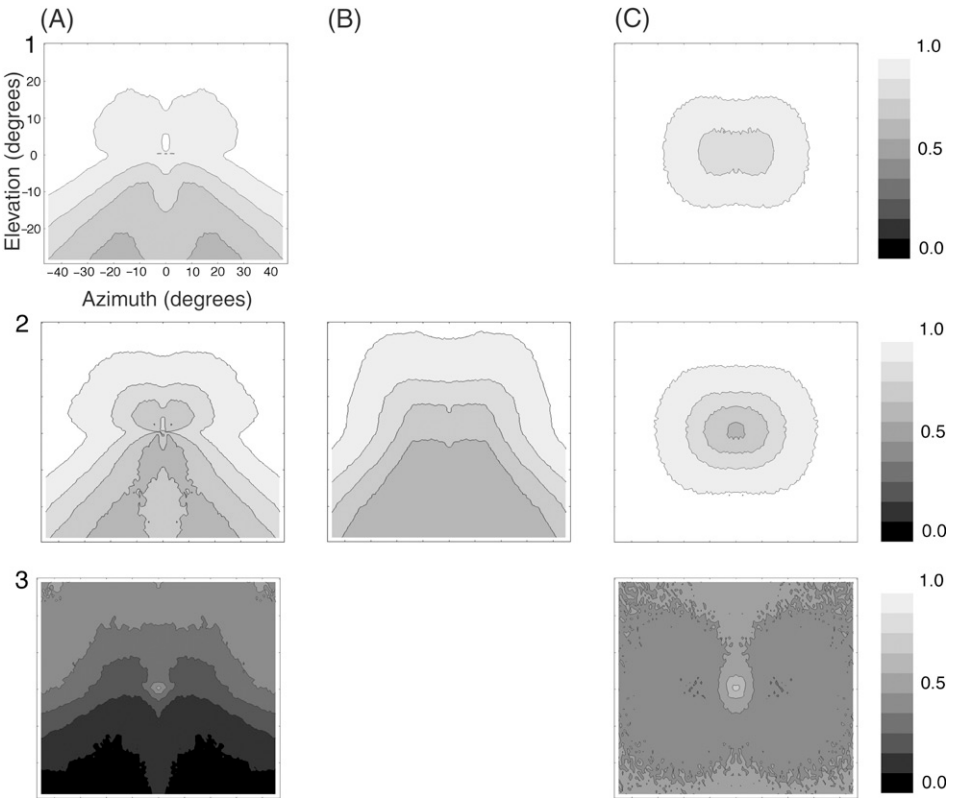

Figure 10. Statistical dependence between speed and depth. Estimated normed mutual information in the different conditions (A) natural, (B) nonstabilized and (C) mixed depth, plotted over the visual field. 1: mI^{normed} (V, ($1/Z$, $1/Z_f$)) between retinal speed V and depth structure ($1/Z$, $1/Z_f$), 2: mI^{normed} (V, $1/Z$) between V and the inverse of depth $1/Z$, 3: mI^{normed} (V, Z/Z_f) between V and the quotient of depth and fixation depth Z/Z_f.

depth statistics have a larger influence in the upper visual field than in the lower visual field. This increased dependence on depth in the upper visual field disturbs the linkage to the heading.

Panels A2 and A3 show how the retinal flow directions are influenced by the vertical and horizontal heading components H_θ and H_ϕ separately. H_ϕ has a strong influence on the statistics of retinal flow direction for eccentric positions on the vertical meridian. H_θ has the highest influence at eccentric positions along the horizontal meridian.

In the nonstabilized condition, retinal direction is completely predicted by the heading direction, as all retinal motion is radially away from the heading point (focus of expansion). This means that the mutual information between the distribution of directions and heading is infinite. However, as in the natural condition the different components of heading have different statistical influence on retinal direction for different domains of the visual field, but the generated bulges of high-normed mutual information values are broader than in the natural condition and do not show the abrupt decrease in the center of the visual field (Figure 11,

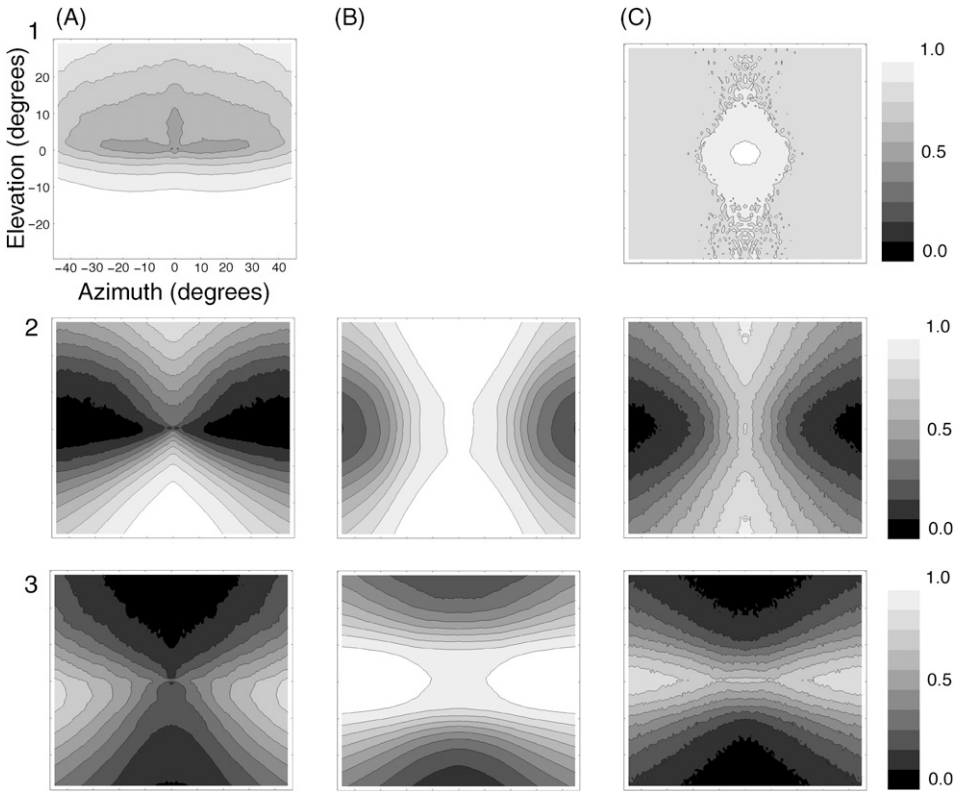

Figure 11. Statistical dependence between direction and heading. Estimated normed mutual information in the different conditions plotted (A) natural, (B) nonstablized and (C) mixed depth, plotted over the visual field. 1: mI^{normed} (Φ_{dir}, (H_ϕ, H_θ)) between retinal direction Φ_{dir} and heading (H_ϕ, H_θ), 2: mI^{normed} (Φ_{dir}, H_ϕ) between Φ_{dir} and the horizontal heading component H_ϕ, 3: mI^{normed} (Φ_{dir}, H_θ) between Φ_{dir} and the vertical heading component H_θ.

Panels B2 and B3). According to the certainty that in the nonstabilized condition the depth structure of the scene has no influence on the statistical behavior of flow directions, there are no differences of the estimated mutual information values between the upper and the lower visual field. In the mixed depth condition the statistical influence of heading on the statistics of flow directions are decreased for the lower visual field and increased for the upper visual field compared with the natural condition (Figure 11, column C). This is a result of the influence of the parameter Z/Z_f on the statistics of retinal flow direction (Figure 10B3). The distribution of normed mutual information for the components H_ϕ and H_θ is similar to that of the natural condition, but rather than drop near the center of view as in the natural condition, the distributions rise in the center of view (Figure 11, Panels C2 and C3).

Figure 12 shows the dependence of the statistics of the retinal speed on the statistics of the observer's heading. In the natural condition (Figure 12A1), the estimated values for $mI^{\text{normed}}(V, (H_\phi, H_\theta))$ (Figure 12A1), which vary between 0 and 0.77, indicate only a modest statistical influence of heading on the statistics of retinal speed. Similar to retinal direction, the influence is more pronounced in the lower visual field than in the upper visual field.

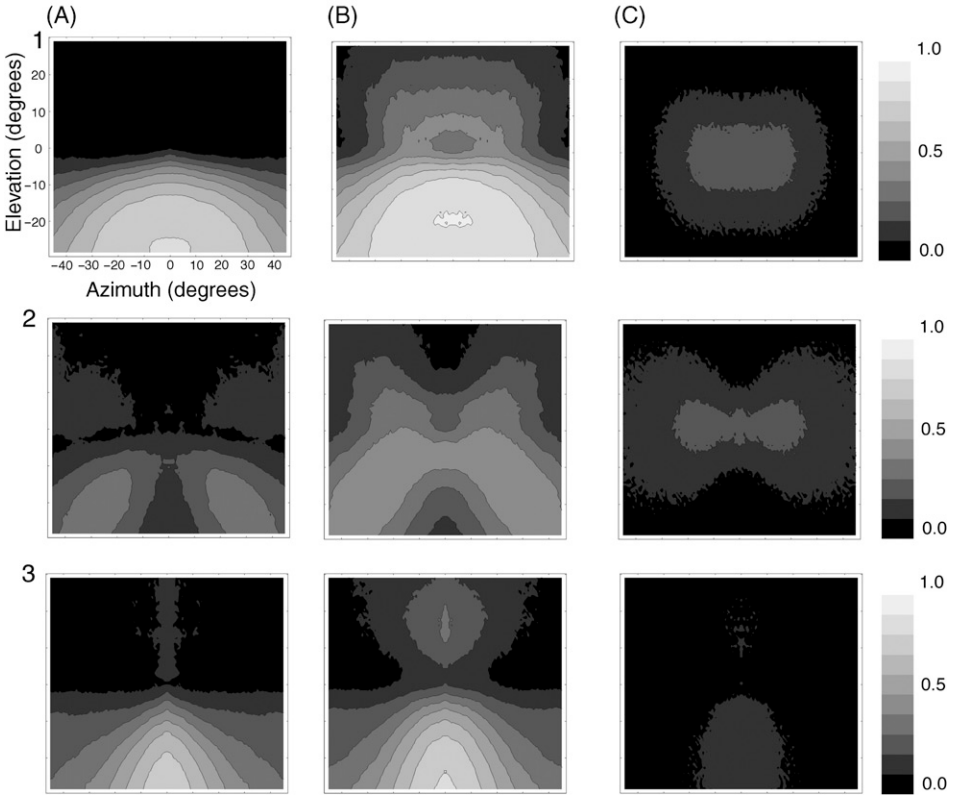

Figure 12. Statistical dependence between speed and heading. Estimated normed mutual information in the different conditions (A) natural, (B) nonstabilized and (C) mixed depth, plotted over the visual field. 1: $mI^{\text{normed}}(V, (H_\phi, H_\theta))$ between retinal speed V and heading (H_ϕ, H_θ), 2: $mI^{\text{normed}}(V, H_\phi)$ between V and the horizontal heading component H_ϕ, 3: $mI^{\text{normed}}(V, H_\theta)$ between V and the vertical heading component H_θ.

The restriction of the statistical influence to the lower visual field is also seen for the separate vertical and horizontal heading components H_θ and H_ϕ (Figure 12A2 and A3). This correlates with the diminished effect of the statistics of depth on the retinal speed for the lower visual field. However, the heading components H_ϕ and H_θ assert their influence in different parts of the lower visual field. H_ϕ influences retinal speed along the diagonals in the lower visual field whereas H_θ influences retinal speed near the vertical meridian in the lower visual field.

In the nonstabilized condition (Figure 12, column B) there is more statistical dependence in the upper visual field but the dependence in the lower visual field is very similar to that in the natural condition. In the mixed depth condition, heading has hardly any statistical influence on retinal speed (Figure 12, column C).

Together, the different influences of H_ϕ and H_θ on retinal direction (Figure 11) and retinal speed (Figure 12) can be explained by the following consideration: a horizontal deviation of heading from straight ahead causes a deviation of the retinal flow direction from radial for the flow elements close to the vertical meridian. Close to the horizontal meridian the direction of flow vectors is less affected, because

mainly the speeds are increased or reduced there. Analogously, the statistical dependence of flow direction along the horizontal meridian is higher for the vertical heading direction. However, the counter-rotation of the retina in the case of gaze stabilization affect the statistical influence of the components of heading by keeping the directions closer to radial.

These observations show that the statistical dependencies of retinal optic flow on heading are shaped by both the geometry of natural scenes and the properties of natural ego-motion. Altering either the ego-motion parameters, or the depth statistics, considerably changes the dependence structure between retinal optic flow and heading.

Summary and discussion

The results show how the structure of the retinal flow depends on the scene statistics and the ego-motion statistics. The principle dependence of the retinal flow on these parameters is clear from the geometrical properties of flow generation (Longuet-Higgins Prazdny 1980). However, the particular relevance of individual parameters in natural situations depends on the statistics of these parameters in the natural context. In the natural condition, the random variables' retinal speed and retinal direction show rather low statistical interdependencies at almost all positions in the visual field. This statistical independence between retinal speed and retinal direction in the natural condition allows to efficiently encode both parameters independently, as is the case in motion-sensitive neurons in visual cortical area MT. These neurons have largely independent tuning curves for direction and speed (Rodman Albright 1987). Gaze stabilization plays an important role for the independence between retinal speed and direction. Without gaze stabilization there is a much higher statistical interdependence between retinal speed and retinal direction for large domains of the lower visual field. Since this increase occurs only in the lower visual field, the depth structure appears to have also a strong influence on these interdependencies.

The distributions of retinal speed and retinal direction are strongly influenced by the underlying statistics of depth and ego-motion parameters. The statistical properties of the distributions measured for the different conditions differ strongly in their behavior over the field of view. In the natural condition, differences between the upper and the lower visual field are clearly visible for both retinal flow direction and retinal speed. In the nonstabilized condition, in contrast, differences between upper and lower visual field only occur in the distributions of retinal speed. This is because depth has no influence on flow direction in the nonstabilized condition. The statistical differences for the upper and the lower visual field are caused by different depth statistics for the upper and the lower visual field. Most natural scenes consist of objects on a ground surface. The ground surface may be flat, or form dips and humps, or it can decline or rise. But in each scene, the ground confines the maximal depth at each positions of the visual field. Therefore, the existence of a ground in natural scenes generates an asymmetry in the depth statistics between positions in the upper and the lower visual field and restricts the variability of depth in the lower visual field. This asymmetry underlies all observed asymmetries between upper and lower visual field in the flow statistics. When the asmmetry in the depth statistics are destroyed in the mixed-depth condition, the differences in optic

flow statistics between the upper and the lower visual field vanish. All statistical variations between different positions in the field of view in that condition are caused by the mathematical rules behind optic flow generation and the statistics of heading. However, the asymmetries between upper and lower visual field in the natural condition arise not just from the depth distribution alone, but rather from a combination of the depth distribution and the natural ego-motion parameters. For instance, many properties of the flow distributions in Figure 11 are symmetric between upper and lower field also in the nonstabilized gaze condition. Regarding the properties of early motion detectors, our results coincide with the findings in Zanker and Zeil (2005), who also state differences in the distributions of the responses of early motion detectors between the upper and lower visual field for straight motion through natural scenes.

One may predict properties of motion-sensitive neurons from the statistics of retinal optic flow according to the principle of efficient encoding, particular with respect to the dependence of tuning properties on the positions of the receptive field. The variation of the distributions of retinal speed and directions over the visual field may explain the variation of properties of neurons encoding different positions in the visual field. Thus, populations of neurons encoding for optic flow near the center of the visual field should be more sensitive for low speed but for a large range of retinal directions, whereas populations of neurons encoding for optic flow more peripherally should be more sensitive for large retinal speed but for more radial retinal directions (cf. for instance (Albright 1989) for such distributions in the primate cortical area MT). The tuning curves of such neurons should account for quantities such as skew and kurtosis, which might affect the proportion of sharply and broadly tuned neurons as well as the tuning in individual cells. Furthermore, the peripheral increase of the size of the receptive field of motion-processing neurons in area MT seems to be well adapted to the structure of natural flow fields (CAlow et al. 2005). More quantitative predictions may be derived from an analysis of efficient encoding of optic flow based on the measured distributions.

Our analysis reveals that the dependence of retinal speed and direction on the set of ego-motion and scene parameters varies considerably across the visual field. In the natural condition, the influence of the depth statistics on the retinal speed is strongest in the upper visual field and much weaker in the lower visual field. Since optic flow depends on depth and heading, an increase in the statistical influence of one parameter must be accompanied by a decrease of the statistical influence of another parameter. Therefore, the reduction of the influence of the depth statistics on the retinal speed coincides with an increased influence of the statistics of heading on the statistics of retinal speed at the lower visual field. This is true also in the nonstabilized condition. The finding that the dependence of retinal speed on depth is highest in the mixed depth condition shows that retinal speed can be regarded as directly linked to the depth map of the scene, at least for the upper visual field. For the lower visual field, the dominance of the ground and the associated decrease in the variation of the depth over different scenes increases the statistical influence of the remaining parameters.

The statistics of retinal direction in the natural condition, on the other hand, are independent of the statistics of depth throughout most of the visual field. The same is true in the mixed depth condition. However, retinal direction is linked to the statistics of the combination of depth and depth of fixation in terms of the

quotient Z/Z_f. This quotient separates the scene in foreground (entities closer than the fixation point, $Z/Z_f < 1$) and background (entities more distanced than the fixation point, $Z/Z_f > 1$). The dependence of direction on Z/Z_f is most pronounced near the horizontal meridian, presumably because variation in depth relative to the depth of fixation occurs most frequently in that area of the scene image.

It is conceivable that the tight statistical linking of optic flow to the depth structure of the scene enables the brain to reconstruct a good relative depth map of the scene from the motion signals. In the natural condition, heading influences the statistics of direction and speed of the retinal motion to different degrees. This is especially pronounced in the lower visual field. Heading has the largest influence on the statistics of direction. The influence of heading on retinal speed remains minor. The azimuth and the elevation component of heading have mutually exclusive statistical influence on the retinal flow direction. The azimuth component and the elevation component of heading are statistically independent ($mI^{\text{normed}}(H_\phi, H_\theta) = 0.0032$) in the distribution we used. The strong statistical influence of one heading component in a certain domain of the visual field leads to a high correlation between the directions of flow vectors within that domain and to a lower correlation between the directions of flow vectors lying in domains which are influenced by the other heading component. We leave the quantitative investigation of the statistical correlation between flow vectors at different position, of the field of view and the extraction of possible independent components or patterns for future work. But detection of such flow patterns requires receptive fields, which fully contain the extend of the pattern. Therefore, the sizes of the domains of a high statistical influence of a certain heading component on retinal flow might also predict the sizes of the receptive fields of heading sensitive neuron. Heading estimation from optic flow is processed in the medial superior temporal (MST) brain area, which receives most of the incoming information from motion-sensitive neurons in area MT and which is widely accepted to process patterns of optic flow (Tanaka and Saito 1989; Duffy and Wurtz 1991). The large receptive field of neurons in area MST are therefore consistent with the large extend of the domains of a high statistical influence of the azimuth and the elevation component of heading on retinal optic flow, particularly in the periphery.

Our study was performed with human-like ego-motion. It is difficult to speculate how the local statistics of retinal velocity will differ for other animals. The height of the eyes above ground may quantitatively alter the statistics, particularly in the lower visual field. However, Zanker and Zeil (2005) reported that the properties of motion signal distributions in the upper visual do not change much with variation of the height of the camera field. Many higher animals that live mostly on the ground, in similar environments, and perform gaze-stabilization reflexes will qualitatively encounter similar local statistics of optical flow. To what extend the statistics will change quantitatively for different species is an interesting question for future work.

Acknowledgements

ML is supported by the German Science Foundation DFG LA-952/2 and LA-952/3, the German Federal Ministry of Education and Research, and the EC Project Drivsco.

References

Albright TD. 1989. Centrifugal directionality bias in the middle temporal visual area (MT) of the macaque. Vis Neurosci 2:177–188.

Atick JJ, Redlich AN. 1992. What does the retina know about natural scenes? Neural Comp 4:196–210.

Barlow HB. 1961. Possible principle underlying the transformation of the security message. In: WA Rosenblith, editor. Sensory communication. Cambridge, MA: MIT Press. pp 217–234.

Berkes P, Wiskott L. 2002. Slow feature analysis yields a rich repertorie of complex cell properties. Proc Int conf on Artifical Netural Network, ICANN02, pp 81–36.

Betsch B, Koerding K, Einhäuser W, König P. 2004. The world from a cats perspective – statistics of natural videos. Biol Cybern 90:41–50.

Calow D, Krüger N, Wörgötter F, Lappe M. 2004. Statistics of the optic flow for self–motion through natural sceneries. In: U Ilg, HH Bülthoff, HA Mallot, editors. Dynamic perception 2004. Berlin: Akademische Verlagsgesellschaft Aka GmbH. pp 133–138.

Callow D, Krüger N, Wörgötter F, Lappe M. 2005. Biologically motivated space–variant filtering for robust optic flow processing. Network comput Neural sys 16(4):323–340.

Duffy CJ, Wurtz RH. 1991. Sensitivity of MST neurons to optic flow stimuli. I. A Continuum of response selectivity to large–field stimuli. J Neurophysiol 65:1329–1345.

Elder EH, Goldberg RG. 2002. Ecological statistics of Gestalt laws for the perceptual organization of contours. J Vision 2:324.

Fermüller C, Shulman D, Aloimonos Y. 2001. The statistics of optical flow. Comp Vis Image Understand. 82:1–32.

Gibson JJ. 1950. The Perception of the Visual World. Boston: Houghton Mifflin.

Gibson JJ. 1966. The senses Considered As Perceptual Systems. Boston: Houghton Mifflin.

Huang J, Lee AB, Mumfort D. 2000. Statistics of range image. Proc CVPR 1:324–321.

Imai T, Moore ST, Raphan T, Cohen B. 2001. Interaction of the body, head, and eyes during walking and turning. Exp Brain Res 136:1–18.

Ivins J, Porill J, Frisby JP, Orban G. 1999. The 'ecological' probability density function for linear optic flow: Implications for neurophysiology. Perception 28:17–32.

Kalkan S, Calow D, Felsberg M, Wörgötter F, Lappe M, Krüger N. 2005. Local image structures and optic flow estimation. Network Comput Neural Sys 16(4):341–356.

Kozachenko LF, Leonenko NN. 1987. Sample estimate of the entropy of a random vector. Probl Peredachi Inf 23:95–101.

Kraskov A, Stgbauer H, Grassberger P. 2004. Estimating mutual information. Phys Rev E 69:06613801–06603816.

Krüger N. 1998. Collinearity and parallelism are statistically second order relation of the complex cell responses. Neural Process 8(2):117–129.

Krüger N, Wörgötter F. 2002. Multi–modal estimation of collinearity and parallelism in natural image sequences. Comput Neural Sys 13:553–576.

Lappe M. (1996). A model of a goal–directed integration of the motion and stereopsis in visual cortex. Preception 25:133.

Lappe M. 2000a. Computational mechnisms for optic flow analysis in primate cortex. In: M Lappe, editor. Neuronal processing of optic flow, Int Rev Neurobiol. Vol. 44. San Diego, London: Academic Press. pp 235–268.

Lappe M. editor. (2000b). Neuronal processing of optic flow, Int Rev Neurobiol. Vol. 44. San Diego, London: Academic Press.

Lappe M, Bermmer F, Pekel M, Thiele A, Hoffmann K. 1996. Optic flow processing in monkey STS: A theoretical and exeperimental approach. J Neurosci 16(19):6265–6285.

Lappe M, Bremmer F, van den Berg AV. 1999. Perception of self-motion from visual flow. Ternds Cogn Sci 3:329–336.

Lappe M, Pekel M, Hoffmann KP. 1997. Optokinetic eye movements elicited by radial optic flow in macaque monkeys. Sco Neurosci Abstr 23:758.

Lappe M, Pekel M, Hoffmann KP. 1998. Optickinetic eye movement eye movements elicited by radial optic flow in the macaque monkey. J Neurophysiol 79:1461–1480.

Laughlin SB. 1981. Simple coding procedure enhances a neuron's information capacity. Z Naturforsch 36C:910–912.

Longuet-Higgins HC, Prazdny K. 1980. The interpretation of a moving retial image. Proc Royal Sco London B 208:385–397.

Niemann T, Lappe M, Bscher A, Hoffmann K-P. 1999. Ocular responses to radial optic flow and singel accelearted targets in humans. Vis Res 39:1359–1371.

Olshausen BA, Field DJ. 1996. Emergence of simple–cell receptive field properties by learning a sparse code for natural image. Nature 381:607–609.

Parzen E. 1962. On extimation of a probility density function and mode. Ann Math Statistic 33:1065–1076.

Rodman HR, Albright TD. 1987. Coding of visual stimulus velocity in area MT of the macaque. Vis Res 27(12):2035–2048.

Roth S, Black MJ. 2005. On the spatial Statics of optical flow. IEEE Int Conf on Comp Vision (ICCV) 1:42–49.

Rudermann DL, Bialek W. 1994. Statistics of natural images: scaling in the woods. Phy Rev Let 39:814–817.

Saito HA, Yukie M, Tanaka K, Hikosaka K, Fukada Y, Iwai E. 1986. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. J Neurosci 6:145–157.

Silverman BW. (1986). Density estimation. London: Chapman and Hall.

Simoncelli EP, Olshausen BA. 2001. Natural image statics and neural representation. Ann Rev Neurosci 24:1193–1216.

Solomon EP, Olshausen BA. 1992. Stabilization of gaze during circular locomotion in light: I. Compensatory head and eye nystagmus in the running monkey. J. Neurophysiol 67:1146–1157.

Tanaka K, Saito H-A. 1989. Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the running monkey. J Neurophysio 62:626–641.

Van Hateren JH, Rudermann DL. 1998. Independent component anaysis of natural image sequences yield spatio-temporal filters similar to simple cells in primary visual cortex. Proc Royal Sco London B 265(1412):2315–2320.

Weiss Y, Fleet D. 2001. Velocity likelihoods in biological and machine vision. In: RPN Rao, BA Olshausen, MS Lewicki, editors. Probabilistic Models of the Brain, Preception and Neural Function. Cambridge, Massachusetts: The MIT Press, Massachusetts Institute of Technology. pp 81–100.

Zanker J, Zeil J. 2005. Movement–induced motion signal distribution in outdoor scenes. Network comput Neural Syst 16(4):357–376.

Zetsche C, Krieger G. 2001. Nonlinear mechanism and higher–order statistics in biological vision and electronic image processing: Review and perspectives. J Electron Imageing 10:56–99.