

Statistics of optic flow for self-motion through natural scenes

Dirk Calow¹, Norbert Krüger³, Florentin Wörgötter² and Markus Lappe¹

1 Dept. of Psychology, Westf.-Wilhelms University, Fliednerstr. 21, 48149 Münster, Germany
calow@psy.uni-muenster.de, mlappe@uni-muenster.de

2 Institute for Neuronal Computational Intelligence and Technology, Stirling, Scotland, UK
worgott@cn.stir.ac.uk

3 Dep. of Computer Science and Engineering, Aalborg University Esbjerg, Denmark
nk@cs.aue.auc.dk

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. We present a method to investigate the statistics of natural optic flow fields related to self-motion within natural scenes and present first results of a statistical analysis.

1 Introduction

In many situations the brain has to analyze ambiguous sensory signals. In such cases it may (re)construct perception by using statistically plausible predictions and/or statistical models of the signal-sending environment. For the sake of efficiency, adaptations of the signal processing system to the statistics of natural environments are very likely. Effects of such adaptations are often seen in Gestalt laws [5]. In the visual modality, several researchers invested effort to reveal the statistics of natural environments and to link them with the neural representation of the sceneries [13, 1, 12, 15, 17, 4, 14, 2]. However, their investigations were largely restricted to static attributes of natural scenes, even when the stimulus material contained dynamic visual scenes [15, 2]. The resulting statistics are treated as invariant with respect to position in the view field. In contrast the properties of motion signals elicited on motion detectors by self-motion within natural sceneries strongly depend on position in the view field. Such motion signals form global patterns of optic flow and there is convincing evidence that these motion patterns are processed in higher visual areas, specifically areas MT and MST [10, 9]. We hypothesize that the motion-processing pathway of the brain may also use statistical properties of natural flow fields to construct optic flow from the motion signals obtained from early motion detectors. An investigation of the statistical properties of optic flow can be the starting point to reveal some interrelations between natural statistics and biological implementations along the motion-processing pathway.

Why is the construction of optic flow from early motion signals important? The pattern of optic flow generated by self-motion encodes behaviorally relevant information about the direction and velocity of self-motion, the distances of potential obstacles, and more [10, 9]. Animals use this information for path planning, obstacle avoidance, ego-motion control and foreground-background segregation to recognize moving objects. Early motion signals are estimated from the spatiotemporal intensity changes of light falling onto the retina and are affected by the aperture problem, noise in the processing pathway, spatial luminance fluctuations, etc. Models of early motion detection have



Figure 1 Panoramic projection of the 3D data of a range-image, brightness encodes the intensity of the reflected laser beam, i.e. white: high intensity, black: vanishing intensity.

demonstrated that these motion signals are ambiguous and often simply wrong [3, 7]. Use of statistical relationships may help to disambiguate and regularize the flow field. Statistical analyses of optic flow may be undertaken on the true motion signals, or the signals from early motion detectors, or the correlation of both. An investigation of the properties of motion fields estimated from elementary motion detectors of the fly can be found in [16]. An analysis of the relation between the flow fields estimated by several flow algorithms and the actual optic flow can be found in [7]. The present investigation is dedicated to the statistics of the true motion signals. For such an analysis a large set of correct optic flow fields obtained from various self-motions through natural environments are required. In the following we introduce a method to generate correct optic flow fields for self-motion through natural sceneries in sufficient number. We then present the first results of a first order statistics of optic flow fields for natural self-motion.

2 Methods

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 [6]. 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 2-200m. The laser wavelength is $0.9\mu\text{m}$ in the near infrared region. Thus the data of each point consist of 4 values, the horizontal angle ϕ , the vertical angle θ and the distance $R(\phi, \theta)$ in spherical coordinates and a value for the reflected intensity of the laser beam. The location of the source of the laser beam is 1.5m above the ground. The 197 data sets can be categorized as "forest", "residential", and "interior" sceneries. Figure 1 shows a typical range-image of the category "residential" projected onto the ϕ - θ plane. Note that the horizon is located at $\theta=90^\circ$. 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.

Retinal projection. The knowledge of the 3D coordinates of each image point allows calculation of the true motion of the corresponding projected 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 a spherical projection surface. To describe optic flow vectors on the sphere we use 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 counterclockwise. Note that 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 of a point moving over the sphere described in terms of the temporal derivations of ϵ and σ is given by $v=(d\epsilon/dt, \sin(\epsilon)d\sigma/dt)$. Thus one obtains

$$\begin{aligned} v_\epsilon &= \cos(\epsilon)\sin(\epsilon)T_z/Z - (\cos(\sigma)T_x + \sin(\sigma)T_y)\cos^2(\epsilon)/Z + \sin(\sigma)\Omega_x - \cos(\sigma)\Omega_y, \\ v_\sigma &= (\sin(\sigma)T_x - \cos(\sigma)T_y)\cos(\epsilon)/Z - \sin(\epsilon)\Omega_z + \cos(\epsilon)(\cos(\sigma)\Omega_x + \sin(\sigma)\Omega_y), \end{aligned}$$

where $T=(T_x, T_y, T_z)$, $\Omega=(\Omega_x, \Omega_y, \Omega_z)$, and $Z=R(\phi, \theta)\sin(\phi)\sin(\theta)=R(\epsilon, \sigma)\cos(\epsilon)$ denote the translation and rotation of the projection surface, and the depth at position (ϵ, σ) .

Ego-motion parameters. To calculate the flow field from the scene structure we need the motion parameters of the projection surface. These involve movement directions within the scenes, and rotations that stabilize gaze on environmental objects as the observer moves through the scene [8, 9, 11]. Because of gaze stabilization reflexes we assume that the rotation depends on the translation by $\Omega=1/Z_f(T_y, -T_x, 0)$. Z_f denotes the depth of the point at which gaze is directed.

To determine possible movement direction within a range image scene we search for areas in the scene, which are free of obstacles in a depth of at least 3m and a width of 0.7m. This criterion gives us a set of possible moving direction for each scene. To obtain gaze directions that we can use to generate gaze stabilization movements we measured eye movements of observers who viewed the scenes on a 17-inch computer monitor. Segments of the range images (white frame in figure 1) that were centered on possible movement directions and projected on a 36.5cm×27.5cm plane with a focal length of 30cm were used for the presentations. Six subjects viewed these images with the head stabilized on a chin rest 30cm in front of the monitor. Gaze fixation points were measured by an eye tracking system (Eye Link II). Pictures were shown for 1 second in immediate succession. The first fixation for each picture was rejected because it might be partially driven from the preceding picture. The subsequent fixations were used as gaze directions for the later analysis. Figure 2 shows gaze directions while viewing "residential" sceneries. The points are plotted in spherical coordinates (ϕ, θ) and centered on the direction of motion used for the flow field calculations.

Optic flow fields. To calculate the flow fields, we assume that the subjects are moving through the sceneries with a velocity of 1m/s in a certain direction as described above. The field of view for each flow field is set to 90° horizontally and 58° vertically. This field of view is subdivided in pixels of 0.36°×0.36° yielding a grid of 250×160 pixels. As the angular separation of the range images is 0.18 one pixel covers up to 4 data points. The depth values Z from these data points are averaged. The mean depth value is

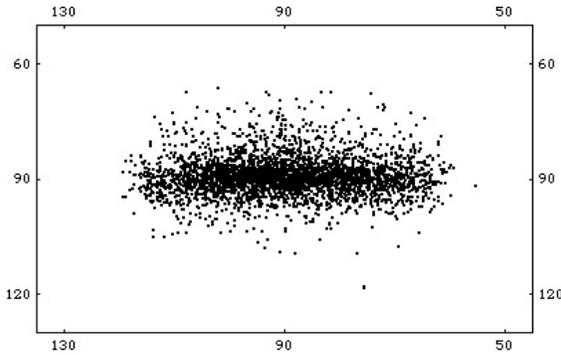


Figure 2: Measured gaze directions plotted in spherical coordinates (ϕ, θ) .

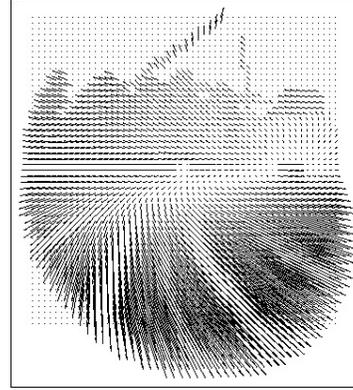


Figure 3: Retinal flow field of the scene in the white frame in figure 1 for a particular ego-motion.

assigned to the pixel in question. The flow vector of a pixel depends on the depth value, the translation and rotation components of the ego-motion and the visual field position of the center of the pixel. The flow vectors of all pixels of an image provide the correct flow field for this gaze direction, ego-motion, and scene. Figure 3 shows an example of a true retinal flow field.

Statistical analysis. In this way, we obtain 2558 flow fields for the "residential" category, 1017 flow fields for the "forest" category and 408 for the "indoor" category. To examine the first order statistics of these flow fields the two dimensional velocity space (v_e, v_o) is subdivided in bins of $0.286479 \text{ degree/s} \times 0.286479 \text{ degree/s}$. Each flow vector falling into a velocity bin increases the frequency of occurrence for the bin by 1. The first order statistics we obtain comprise the relative frequencies in velocity space as a function of the position in the visual field.

3 Results

Figure 4, left panel, shows the frequency distributions of velocities as density plots for six positions in the left visual hemifield. The distributions in the right hemifield are essentially mirror-symmetric. To better visualize the distributions, the grey scale shows the third root of the relative frequencies. Thus the maximum value 0.19 pictured as black points is equivalent to a relative frequency of 0.0068. Note that the horizontal meridian is defined by $\theta = 90^\circ$ and the vertical meridian by $\phi = 90^\circ$.

Figure 4 (left panel) reveals clear differences between the upper and the lower field of view. In the lower field (here presented by the positions $\phi = 126^\circ, \theta = 112^\circ$ and $\phi = 90^\circ, \theta = 112^\circ$) the distributions are rather broad and spread out in the positive v_e half space. At $\phi = 90^\circ, \theta = 112^\circ$ there is a symmetric deviation from the radial direction in the positive and negative v_o direction. For position $\phi = 126^\circ, \theta = 112^\circ$ this distribution is asymmetric. There is a correlation between radial velocity v_e and tangential velocity v_o such that for smaller v_e values ($v_e < 0.2 \text{ rad/s}$) the distribution is spread in the positive v_o half space and

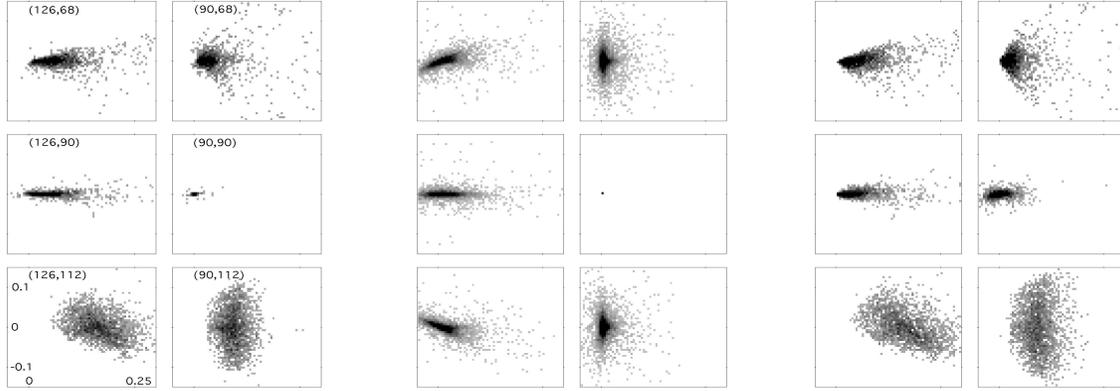


Figure 4: Relative frequencies for 6 different view field positions. Left: natural scenes, middle: random scenes, right: natural scenes without gaze stabilization. The axis's are labeled in rad/s. Positions: Top left: $\phi=126^\circ$, $\theta=68^\circ$, top right: $\phi=90^\circ$, $\theta=68^\circ$, middle left: $\phi=126^\circ$, $\theta=90^\circ$, middle right: $\phi=90^\circ$, $\theta=90^\circ$, down left: $\phi=126^\circ$, $\theta=112^\circ$, down right: $\phi=90^\circ$, $\theta=112^\circ$

conversely for higher v_ϵ values in the negative v_σ half space.

In general, caused by the smaller depth values in the lower visual field, the velocity values are higher and the distribution is broader than in the upper visual field. Since the depth values in the upper visual field are generally larger the distributions are less broad and values of $v_\epsilon > 0.17 \text{ rad/s}$ are rare. But like in the lower view field the distributions are skewed towards the positive v_ϵ half space. At $\phi=126^\circ$, $\theta=90^\circ$, i.e. along the horizontal meridian, the velocities are largely radial. Because any deviations from the radial direction at this position can only be caused by the rotation component Ω_x from the gaze stabilization, this distribution reflects the effects of the largely horizontal distribution of gaze positions (Figure 2).

It is interesting to see to what degree the properties of these distributions depend on the structure of the environment or on the ego-motion parameters. To reveal the influence of scene structure we compare the respective distributions (left panel in figure 4) with frequency distributions obtained with the same ego-motion parameters in a completely unstructured environment (middle panel in figure 4). In this unstructured environment depth is random and uniformly distributed between 2m and 50m. In the unstructured environment the distributions are symmetric between the upper and lower visual field of view. The kernels (black regions) have the shape of slightly distorted triangles and are more peaked than in the natural environment. Furthermore the distributions in natural environments are shifted towards higher v_ϵ values.

To reveal the influence of the ego-motion parameters, particularly gaze direction and gaze stabilization movements, we compare the distributions of natural scenes (left panel in figure 4) with frequency distributions obtained with the same scenes but without gaze stabilization (right panel in figure 4). It can be seen that respective distributions become broader without gaze stabilization. Especially at position $\phi=90^\circ$, $\theta=90^\circ$ the missing gaze stabilization leads to motion signals also in the central field of view.

4 Conclusions

The first order statistics of retinal optic flow reveal a clear influence of both the structure of natural scenes and the ego-motion parameters. Large differences between natural and completely unstructured environments show the influence of scene statistics. The existence of the ground in natural environments leads to clear differences between the optic flow statistics of the upper and the lower field of view. Differences between ego-motion with and without gaze stabilization show the importance of gaze stabilization reflexes. Because of gaze stabilization, flow fields become more stereotyped and the velocity distributions become sharper.

Acknowledgements: ML is supported by the German Science Foundation, the German Federal Ministry of Education and Research BioFuture Prize, the Human Frontier Science Program, and the EC Projects ECoVision and Eurokinesis.

References

- [1] J. J. Atick and A. N. Redlich. What does the retina know about natural scenes? *Neural Comp.*, 4:196–210, 1992.
- [2] K. K. B. Betsch, W. Einhäuser and P. König. The world from a cats perspective - statistics of natural videos. *Biol.Cybern.*, 90:41– 50, 2004.
- [3] J. E. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *Int.J.Computer Vision*, 12:43–77, 1994.
- [4] P. Berkes and L. Wiskott. Slow feature analysis yields a rich repertoire of complex cell properties. *Proc. Int. Conf on Artificial Neural Networks, ICANN02*, pages 81– 86, 2002.
- [5] E. H. Elder and R. G. Goldberg. Ecological statistics of Gestalt laws for the perceptual organization of contours. *J.Vision*, 2:324–353, 2002.
- [6] J. Huang, A. B. Lee, and D. Mumford. Statistics of range images. *CVPR 2000*.
- [7] S. Kalkan, D. Calow, M. Felsberg, F. Wörgötter, M. Lappe, and N. Krüger. Optic flow statistics and intrinsic dimensionality. In *Brain Inspired Cognitive Systems 2004*, 2004.
- [8] T. Niemann, M. Lappe, A. Büscher, and K.-P. Hoffmann. Ocular responses to radial optic flow and single accelerated targets in humans. *Vision Res.* 39, 1359-71, 1999.
- [9] M. Lappe, editor. *Neuronal Processing of Optic Flow*, Academic Press, 2000.
- [10] M. Lappe, F. Bremmer, and A. V. van den Berg. Perception of self-motion from visual flow. *Trends.Cogn.Sci.*, 3:329–336, 1999.
- [11] M. Lappe, M. Pekel, and K.-P. Hoffmann. Optokinetic eye movements elicited by radial optic flow in the macaque monkey. *J.Neurophysiol.*, 79:1461–1480, 1998.
- [12] B. A. Olshausen and D. J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996.
- [13] D. L. Rudermann and W. Bialek. Statistics of natural images: scaling in the woods. *Phys.Rev.Let.*, 39:814–817, 1994.
- [14] E. P. Simoncelli and B. A. Olshausen. Natural image statistics and neural representation. *Ann.Rev.Neurosci.*, 24:1193–1216, 2001.
- [15] J. H. van Hateren and D. L. Rudermann. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proc.Royal.Soc.London B* 265:2315-2320, 1998.
- [16] J. M. Zanker and J. Zeil. An analysis of motion signal distributions generated by locomotion in a natural environment. In R. P. Würtz and M. Lappe, editors, *Dynamic Perception, PAI Proceedings in Artificial Intelligence*. Infix Verlag, St. Augustin, 2002.
- [17] C. Zetsche and G. Krieger. Nonlinear mechanism and higher-order statistics in biological vision and electronic image processing: review and perspectives. *J.Electronic Imaging*, 10:56–99, 2001.