An Efficient Encoding Scheme for Dynamic Visual Input Based on the Statistics of Natural Optic Flow

Dirk Calow and Markus Lappe

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

Abstract. Statistically efficient processing schemes focus the resources of a signal processing system on the range of statistically probable signals. Relying on the statistical properties of retinal motion signals during ego-motion we propose a nonlinear processing scheme for retinal flow. It maximizes the mutual information between the visual input and its neural representation and distributes the processing load uniformly over the neural resources. We derive predictions for the receptive fields of motion sensitive neurons in the velocity space. The properties of the receptive fields are tightly connected to their position in the visual field and to their preferred retinal velocity. The velocity tuning properties show characteristics of properties of neurons in the middle temporal area of the primate brain.

Although the processing power of the brain is huge compared with contemporary artificial signal processing systems, the range of signals a single neuron can process is limited. The visual pathways of the brain show adaptations to the statistics of the natural environment for an efficient processing of the set of signals that the environment provides. Such adaptations are seen in gestalt laws [4, 6] and in efficient encoding schemes [2, 9], in which the processing pathway is more sensitive for signals that occur very frequently than to signals that are very unlikely to occur. However, in many natural situations the visual input is dynamic because animals move. We aim here to apply the concept of efficient encoding to the realm of motion processing to find properties of motion sensitive neurons, that efficiently encode the set of motion signals generated on the retina by natural behavior in natural environments. Most of the retinal motion in natural situations is generated by ego-motion of the observer. Therefore, we concentrate our investigation on retinal motion signals during ego-motion.

1 Statistical properties of natural retinal flow

Our analysis starts from an analysis of the statistics of motion signals generated by ego-motion. We use a method introduced in [3] to generate a large number of naturalistic retinal motion fields from range images of natural scenes and

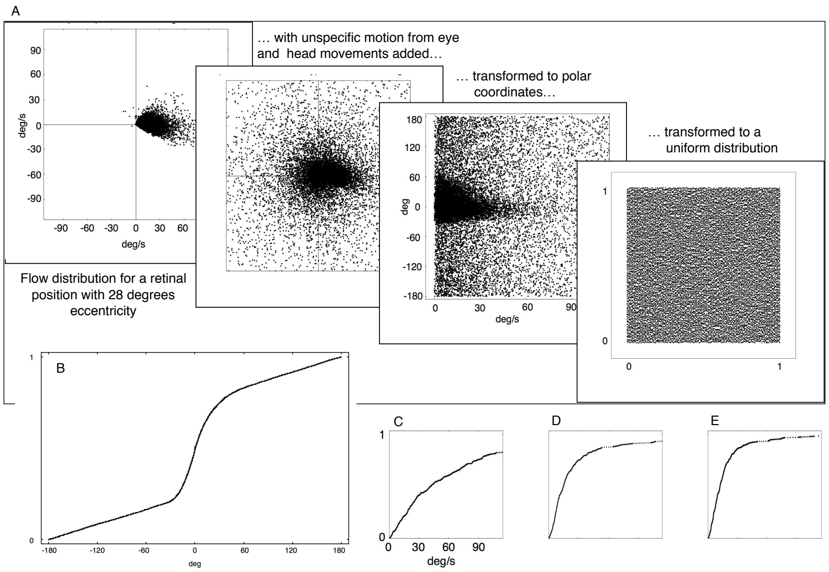

Fig. 1. Uniforming procedure, B: approximation of the transformation function $f_2(\phi)$, C,D and E: approximation of the transformation functions $f_1(v, -180°)$, $f_1(v, -30°)$ and $f_1(v, 0°)$

assumed ego-motions. This approach provides the distributions of true motion signals at each position of the visual field, which then a population of neurons is meant to encode. We use the Brown Range Image Database, a database of 197 range images of outdoor and indoor scenes recorded with a laser range-finder with high spatial resolution [5]. These range images provide the depth map of each scene. The knowledge of the 3D coordinates of each scene point allows the calculation of the projected retinal motion of that point for any given combination of translation and rotation of the observer. We consider as the retina a spherical projection surface with radius 1. The field of view is set to 90° horizontally and 58° vertically. The retinal flow fields are calculated on the inside of this section of the sphere on a 250 × 160 grid of motion sensors with a resolution of 0.36° × 0.36°. The flow vectors obtained for this grid provide our measurement of the true retinal flow field for a certain ego-motion and scene. To calculate the flow field from the scene structure we need the motion parameters of the projection surface. Possible ego-motions for a given scene are derived from properties of the scene and natural parameters of ego-motion. First, we determine areas in each scene which are free from obstacles. We then assumed ego-motion through those areas. Since natural ego-motion involves gaze-stabilizing eye movements [8] we measure eye movements of observers who view images of the scene to collect naturalistic gaze points. We also measure gait dynamics of walking human subjects, particularly their bounce and sway, to allow a naturalistic modeling of the ego-motion trajecory. From these factors (obstacle-free walking direction, gaze point, gait dynamics) we construct a set of naturalistic ego-motions for each range mage scene. From the scenes and the ego-motions we constructed 7136 different naturalistic flow fields and obtain for each position on the retina a distribution of true motion signals (see Figure 1 A, left panel). These flow fields serve to estimate the statistical properties of retinal flow.

2 Efficient encoding of natural retinal flow

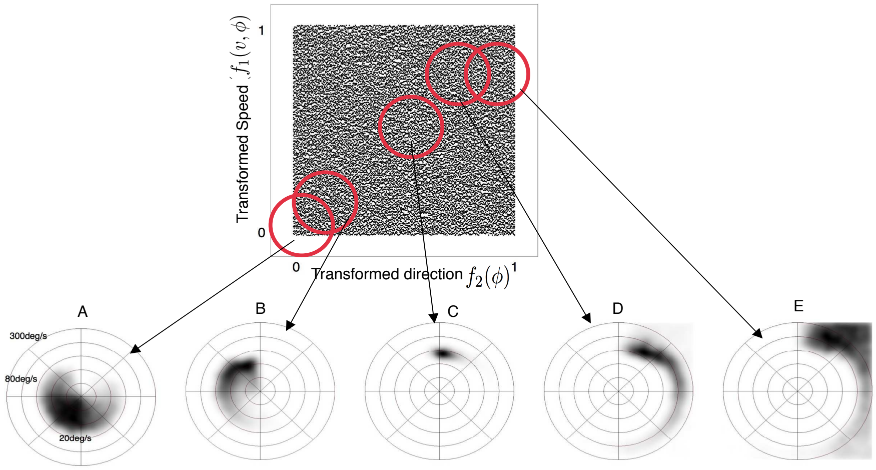

Fig. 2. Resulting responsiveness $r_i(v,\phi)$ of a selection of neurons for certain receptive field centers m_i and width $\sigma = 0.1$ at a retinal position of 28 degrees eccentricity

To turn to our aim of finding an efficient encoding scheme we assume a population of neurons that cover a four-dimensional parameter space consisting of retinal position (azimuth and elevation) and velocity (retinal direction (ϕ_i) and speed (v_i)). This population of neurons efficiently encodes the natural distribution of motion signals when the response probability is the same for all neurons of the population. In this case, the information about the original distribution is uniformly distributed over the neuron population. This maximizes the mutual information between the distribution of motion signals and the distribution of their representation in the neural population, and performs a nonlinear independent component analysis on the original distributions of motion signals. To construct such a population of neurons we search for the transformation functions that maps the distributions of retinal velocities onto a uniform distribution and then cover this uniform distribution uniformly with circular receptive fields (see Figure 1). The actual sensitivity of each neurons for retinal motion signals can then be estimated by the back transformation of the uniform distribution to the original distribution of motion signals. A crucial issue in this procedure is the noise in the signal since the noise level determines the number of neurons necessary to cover the distribution of motion signals and to discriminate between neighboring signals. To each measured distribution therefore we add a set of unspecific motion signals, which is uniform in the distributions of directions and obey a logarithmic normal distribution for retinal speed (see Figure 1 A, second panel) . To find a mapping from the distributions of retinal velocities to a uniform distribution we use a uniforming procedure referred to as rank ordering. The procedure starts with the rank ordering of the distribution of directions, which results in the approximated transformation function for retinal direction $f_2(\phi)$ (see Figure 1 B). The rank ordered data set is then divided into 72 stripes of equal width such that each stripe contains the same number of data points. Regarding each stripe as a one-dimensional data set, the stripes are rank ordered again to yield a set of approximated transformation functions $f_1(v, \phi)$ for retinal speed (see Figure 1 C, D and E). This procedure is an approximation of finding the transformation function $F(v, \phi) = (f_1(v, \phi), f_2(\phi))$ that fulfills $\frac{\partial f_2}{\partial \phi} \frac{\partial f_1}{\partial v} = P(v, \phi)$, where $P(v, \phi)$ is the probability density function of the random variables speed and direction.

3 Tuning curves of efficient retinal flow encoders

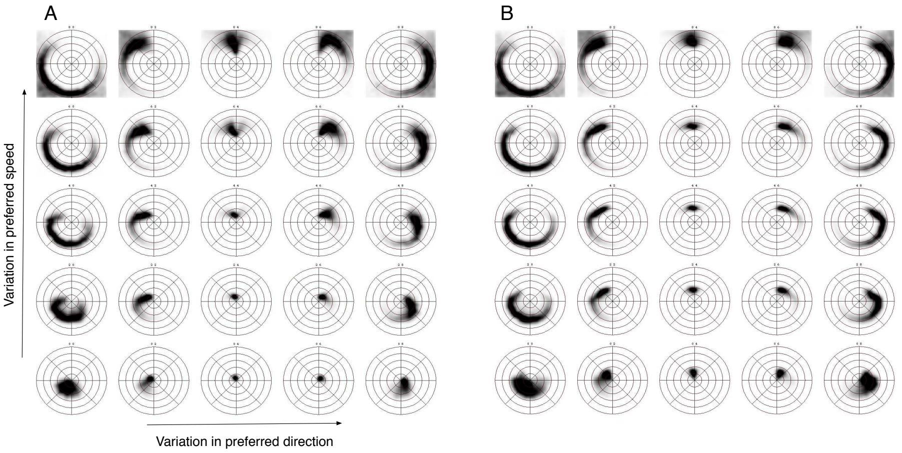

Fig. 3. Resulting responsiveness $r_i(v, \phi)$ of a selection of neurons for certain receptive field centers m_i and width $\sigma = 0.1$ for two different positions in visual field, A: eccentricity 3 degrees, B: eccentricity 15 degrees

The approximation of the transformation functions $F(v, \phi)$ by the uniforming procedure allows to plot the receptive fields in velocity space of neurons which are supposed to encode the distribution of motion signals at a certain position in the field of view. To find the tuning curve of a neuron we have to back-transform from the uniformed space to retinal motion space. Let (v, ϕ) be a retinal motion signal. The mean response $r_i(v, \phi)$ of the i-th neuron to a retinal motion signal is given by $r_i(v, \phi) = r_{max} \exp(-(m_i - F(v, \phi))^2/2\sigma_i^2)$, where $m_i \in [0, 1] \times [0, 1]$ is the center of the receptive field, σ is the width of the receptive field, and $F : \mathbb{R}^2 \longrightarrow [0, 1] \times [0, 1]$, $F(v, \phi) = (f_1(v, \phi), f_2(\phi))$ is the transformation function, which governs the tuning curve of the neuron and which is approximated by the results of the uniforming procedure. The results for a number of example neurons with different receptive field centers m_i are shown in Figure 2 and Figure 3 . The resulting properties of efficiently encoding neurons show a wide range of selectivity for the parameters retinal speed and direction. We find neurons, whose selectivity is largely restricted to a small domain in the velocity space (Figure 2 C). There are also neurons whose selectivity show a crescent-shaped structure (Figure 2 B and D). Such structures resemble receptive field properties of motion processing neurons in the middle temporal (MT) area of the primate [10]. Other similarities to properties of MT neurons can be seen in the speed selectivity. Individual neurons can be characterized as low pass (Figure 2 A), tuned (Figure 2 C), broad band (Figure 2 B and D), or high pass (Figure 2 E) consistent with neurophysiological recordings [7]. With respect to direction tuning, neurons can be sharply tuned, broadly tuned, or not tuned also similar to neurons in area MT. However, the particular properties of the velocity receptive fields of efficiently encoding neurons depend strongly on the position of their receptive field in the visual field (Figure 3). Since the properties of the distributions of retinal motion signals vary from position to position [3], also the population of neurons encoding motion at these positions show differences in their properties. At higher eccentricities, the selectivity for retinal speed shifts to higher speeds and the distribution of preferred directions becomes narrower. This is also similar to findings in area MT [1, 7].

We conclude that the application of the principle of efficient encoding to the processing of retinal motion signals is a valid tool to predict receptive field structures and tuning curves of neurons in the motion processing pathway of the brain. Several interesting issues remain for future work. First, the incfluence of internal noise may be investigated using a Poisson model to describe the response behavior of the neurons. Second, a Bayesian decoding scheme may be applied to develop computational models that detect retinal motion in the statistically likely range. Third, it may be tested experimentally whether humans show a similar characteristics in motion detection.

References

1. T. D. Albright. Centrifugal directionality bias in the middle temporal visual area (MT) of the macaque. *Vis.Neurosci.*, 2:177–188, 1989.
2. H. B. Barlow. Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith, editor, *Sensory Communication*, pages 217–234. MIT Press, Cambridge, MA, 1961.

3. D. Calow, N. Krüger, F. Wörgötter, and M. Lappe. Statistics of optic flow for self-motion through natural sceneries. In U.Ilg, H. H. Bülthoff, and H. A. Mallot, editors, *Dynamic Perception 2004*, pages 133–138, 2004.
4. E. H. Elder and R. G. Goldberg. Ecological statistics of Gestalt laws for the perceptual organization of contours. *J. Vision*, 2:324–353, 2002.
5. J. Huang, A. B. Lee, and D. Mumford. Statistics of range images. *Proc. CVPR*, 1:324–321, 2000.
6. N. Krüger and F. Wörgötter. Multi-modal estimation of collinearity and parallelism in natural image sequences. *Network: Computation in Neural Systems*, 13:553–576, 2002.
7. L. Lagae, S. Raiguel, and G. A. Orban. Speed and direction selectivity of macaque middle temporal neurons. *J.Neurophysiol.*, 69:19–39, 1993.
8. 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.
9. S. B. Laughlin. Simple coding procedure enhances a neuron's information capacity. *Z.Naturforsch.*, 36C:910–912, 1981.
10. M. S. Livingstone, C. C. Pack, and R. T. Born. Two-dimensional substructure of MT receptive fields. *Neuron*, 30:781–793, 2001.