ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 319-324

A Neural Model of Heading Detection from Optic Flow

Frank Seifart , Pierre Bayerl, Heiko Neumann
University of Ulm, Department of Neural Information Processing,

89069 Ulm, Germany

Abstract. This paper describes a neural model developed for computing heading from optic flow caused by 3D translational egomotion. The model uses the distributed representation of optic flow directions in cortical areas MT and MSTd. Model MSTd cells are selective for specific directions of visual motion and have large receptive fields covering approximately a quarter of the visual field at different retinal positions. The estimation of heading is computed in a polar framework by combining the activation of all MSTd cells in a geometrically motivated and biological plausible manner. In our implementation optic flow fields were generated from motion of a simulated camera in a static environment. We analysed the detection error by comparing estimated heading with the ground truth defined by the given camera motion. The results show that the described neural approach provides a robust detection method. We demonstrate that movements inducing radial flow patterns (forward movements) are detected more accurately than motions inducing laminar flow fields (e.g. sideward movements), consistent with psychophysical findings. Most important is that the described properties are a consequence of simple geometrical constraints defined by the spatial arrangement of MSTd cells.

1 Introduction

The determination of the direction of egomotion (heading) is one of the basic tasks of navigation of humans and all higher animals. Self-motion induces typical patterns of optic flow on the retina which are analysed by the visual system. One invariant characteristic of the 2D optic flow is the existence of a ’Focus–of–expansion’ (FoE), a spatial singularity representing the source of expanding flow vectors during translational egomotion. The two successive cortical areas MT (medial temporal area) and MSTd (dorsal medial superior temporal area) are concerned with the processing of optic flow.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 319-324

Visual information in primary cortex (V1) is coded in a space-variant way using a log-polar representation [10]. This representation is substantial for the perception of egomotion. It maps radial, circular and spiral patterns to linear patterns and therefore implies an advantageous encoding of motion patterns induced by forward motions.

Cells in area MT represent proper features of optic flow information , e.g. motion direction and speed, in relatively small regions of the visual field. Area MT contains direction-selective cells with medium-size receptive fields [1], [9]. MT projects to area MSTd where cells have huge receptive fields and are tuned to specifc directions in polar space [5], [6].

Previous models utilize specially adapted receptive fields [2] or extract heading from global optic flow patterns using a learning algorithm [8]. We present an approach in which the structure of MT / MSTd cell connections plays a key role to build up the perception of heading. We assume that cells in area MSTd selective for a specific direction sum over MT cells selective for the same polar motion direction as proposed in the model of MSTd in [7]. The advantages of the polar mapping mentioned above also affect higher cortical areas as MT and MSTd.

Within the large receptive field (Fig. 1, *left*) of an MSTd cell the optic flow vectors resulting from a particular egomotion have very similar directions. For each location of a receptive field, one of n possible preferred directions is selected using a winner–takes–all mechanism. We suggest a specific map-like representation to encode the observers 'heading space'. This map is defined by the v_x and v_y dimensions in which same motion directions from maximum MSTd cell responses form regions shaped like cake pieces. Fig. 1 (*right*) shows the areas of same cell responses (main direction of the flow stimulus) for the cell position of Fig. 1 (*left*).

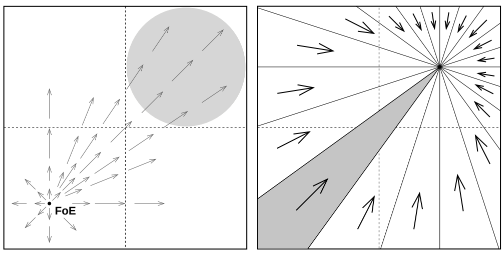

Figure 1: *Left:* Retinal representation of a circular receptive field of a MSTd cell (gray) in the upper right part of the visual field. Arrows indicate the direction of the optic flow stimulus that results from a egomotion to bottom left. *Right:* Areas of the same (discrete) main direction of optic flow. The arrows indicate the main direction. The gray colored area shows the 'activated area' by the flow stimulus from the left side.

Our model takes advantage of these simple geometric properties in order to determine heading by combining the responses of several MSTd cells with receptive fields at different locations in the visual field.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 319-324

2 Neural Model of Heading Detection

Our model aims to detect the direction of translational egomotion from optic flow. The input data (vector field of optic flow directions) is generated by a simulation of a camera moving towards a frontoparallel textured wall. The velocity component in direction of gaze (v_z) was fixed to 1 m / s. The components v_x and v_y form the heading space. To detect heading amounts to determine a certain position in the suggested heading space.

We used a simplified representation of the cortical area MT as an input layer that is composed of cells with spatially medium–size receptive fields. Each model MT cell has a Gaussian receptive field that is tuned around a preferred motion direction [1].

Model area MSTd in turn consists of cells with large receptive fields as shown in Fig. 2. Each MSTd cell receives input from MT cells within a segment covering approximately a quarter of the visual field. A model MSTd cell is selective for a particular preferred direction of optic flow. It is assumed that a segment in model MSTd contains cells selective for any possible preferred direction. The response (main direction) of a set of model MSTd cells located in the same segment is indicated by the preferred direction of the most active cell within the segment (winner–takes–all).

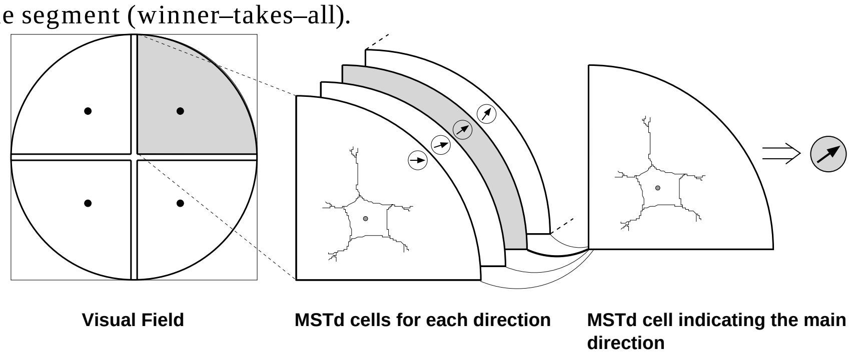

Figure 2: Simplified model of the MSTd area: an example of 4 MSTd segments, each containing MSTd cells selective for any preferred direction. The preferred direction of the most active cell is considered as the segment's response (main direction) using a winner–takes–all method.

As mentioned above, for each MSTd segment we can identify spatial segments like cake pieces in heading space (see Fig. 1) corresponding to FoE's whose expanding flow vectors generate maximum MSTd cell responses. The shape and position of these cake segments can be computed for each model MSTd segment and stored in a lookup table. This is done in the '*training phase*' of our model.

The key idea of the proposed model of heading detection is to combine the responses of MSTd cells located at different retinal positions in visual field. We assume a cell map in higher cortical areas representing the heading space, integrating activation from MSTd. Cells in area 7a were found to show a more

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 319-324

abstract representation of optic flow, which may partly encode the direction of motion [11]. Only recently, area VIP was shown to contain cells that were tuned to heading [3].

Each MSTd cell activates an area (cake segment, computed in the *training phase*) in heading space. We determine heading by adding these activated areas (Fig. 3). The overlap of such regions (Fig. 4, *left*, *center*) leads to regions of maximal activation in heading space, representing the possible directions of egomotion (Fig. 4, *right*). We define the final heading estimation by the median of v_x and v_y values within such regions. This phase is called '*detection phase*'. Thus, heading detection amounts to determine a position of maximum activity, similar to voting spaces such as the the Hough Transform.

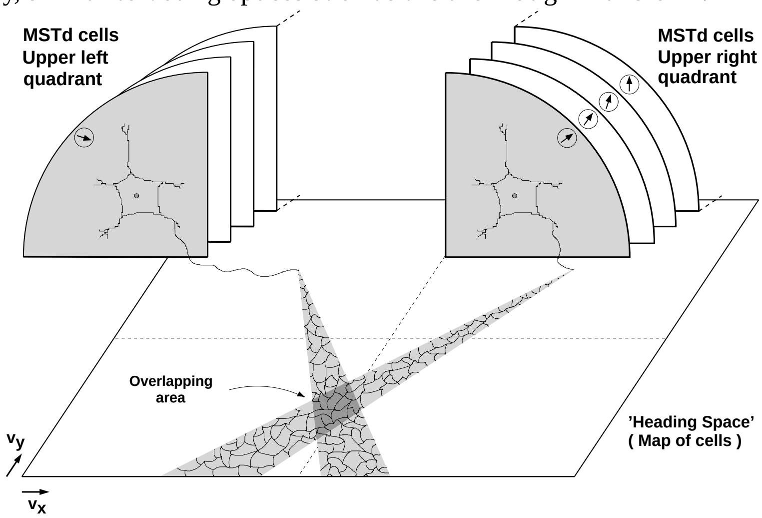

Figure 3: Detection method: Overlapping of activation areas on a cell map representing the heading space

The *training phase* corresponds to the development of connections between direction selective neurons in MSTd area and the neurons of the cell map representing the heading space.

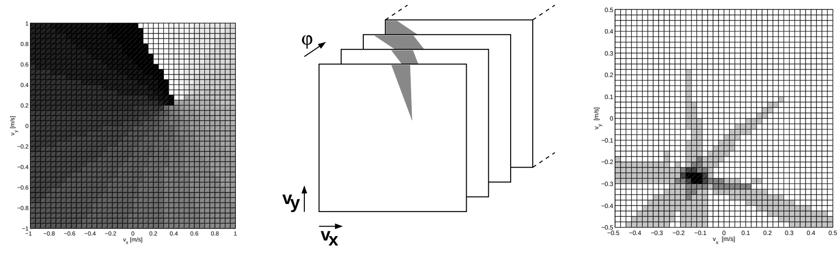

Figure 4: *Left:* Activation areas of a single model MSTd segment shown as grayscale levels. Each gray level represents a preferred direction in heading space. *Center:* Lookup table for this segment. *Right:* Overlapping of activation areas of 4 MSTd segments. Heading is computed by the median of v_x and v_y values from the area of maximum activity (darkest region).

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 319-324

3 Results

To test the proposed detection method we computed lookup tables in the *training phase* for various parameters of the model such as resolution of velocity in heading space, number of possible preferred directions of MT/MSTd cells and position, size and number of MSTd segments. For each configuration the estimated heading is compared with input camera motion direction (ground truth) and the detection error is analysed.

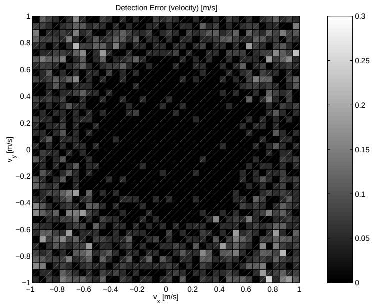

Figure 5: Detection error of an example configuration: 4 MSTd segments (each in one quadrant of the visual field), 40 possible preferred directions, velocity range -1 m/s $\leq v_x/v_y \leq$ 1 m/s, velocity resolution in test motions and in heading space 0.05 m/s. The detection error is coded as grayscale level. The detection error increases towards peripheral regions.

All results show very precise detection performance in the center of heading space (corresponding to radial flow patterns) with an increasing detection error towards peripheral regions (laminar flow patterns) (see Fig. 5). These results are consistent with psychophysical findings from analysing the accuracy of heading perception [4]. The detection error is measured by the absolute deviation of locations in heading space between detected (v_x, v_y) tuples and ground truth.

Further selected results are:

- The detection error decreases with an increasing number of preferred directions (higher number of direction selective cells per segment).
- The detection error depends on the location of areas in the visual field covered by MSTd segments.
- The precision of detection increases if more widely overlapping MSTd segments are used for detection.

We also tested the robustness of the detection method by adding Gaussian noise with standard deviations from 1° up to 30° to the input optic flow directions. The detection error increases with the intensity of noise (Fig. 6). The method, however, still achieves acceptable results under these circumstances. This indicates that the architecture can be used for navigational tasks in real world environments which we currently investigate.

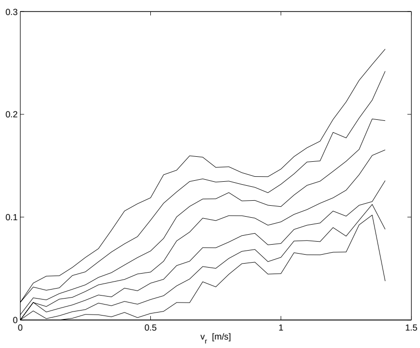

Figure 6: Detection errror for noisy input flow from $\sigma = 0^\circ$ (bottom graph) up to 30° (top graph) and $\Delta\sigma = 5^\circ$.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 319-324

4 Discussion

We developed a neural model for detection of heading from optic flow. Heading is detected by combining the responses of MSTd cells with receptive fields at different locations in the visual field. We found that the model's detection accuracy is best for most common directions of egomotion (forward movements).

The model in its current state of development utilizes some simplified computational principles that can be improved by further developments. In particular, the strongest simplification of the detection method is the winner–takes–all mechanism to select only the maximum response in model MSTd. Storing *all direction responses* in the training phase and using it for overlapping in the detection phase would result in higher computational effort but we suppose an improvement in detection precision. It would also biologically be more plausible. Furthermore, increasing the number of (spatially overlapping) MSTd segments reduces the detection error.

A log-polar compression of the heading space can reduce the amount of memory required by the lookup table. If still the same amount of memory is devoted, a higher detection precision in inner regions of heading space could be achieved.

References

[1] T. D. Albright. Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, 52:1106–1130, 1984.

[2] J. A. Beintema, A. V. van den Berg, and M. Lappe. Receptive field structure of flow detectors for heading perception. In T. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems 14*. MIT Press, 2001.

[3] F. Bremmer, J. Duhamel, S. Ben Hamed, and W. Graf. Heading encoding in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience*, 16(8):1554–1568, 2002.

[4] J. Crowell and M. Banks. Perceiving heading with different retinal regions and types of optic flow. *Perception and Psychophysics*, 53(3):325–337, 1993.

[5] C. J. Duffy and R. H. Wurtz. Sensitivity of MST neurons to optic flow stimuli. I. a continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, 65:1329–1345, 1991.

[6] M. S. A. Graziano, R. A. Andersen, and R. J. Snowden. Tuning of MST neurons to spiral motions. *Journal of Neuroscience*, 14:54–67, 1994.

[7] S. Grossberg, E. Mingolla, and C. Pack. A neural model of motion processing and visual navigation by cortical area MST. *Cerebral Cortex*, 9:878–895, 1999.

[8] M. Lappe, F. Bremmer, M. Pekel, A. Thiele, and K.-P. Hoffmann. Optic flow processing in monkey STS: A theoretical and experimental approach. *Journal of Neuroscience*, 16:6265–6285, 1996.

[9] J. H. R. Maunsell and D. C. van Essen. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, 49:1127–1147, 1983.

[10] E. L. Schwartz. Spatial mapping in the primate sensory projection: analytic structure and relevance to perception. *Biological Cybernetics*, 25:181–194, 1977.

[11] R. Siegel and H. Read. Analysis of optic flow in the monkey parietal area 7a. *Cerebral Cortex*, 7:327–346, 1997.