Dynamical retino-cortical mapping

Markus A. Dahlem and Florentin Wörgötter

Computational Neuroscience
Department of Psychology
University of Stirling
Stirling FK9 4LA
Scotland / UK
{mad1, faw1}@cn.stir.ac.uk
http://www.cn.stir.ac.uk/

Abstract. A dynamical mapping strategy is introduced, that leads to a novel representation of optical flow in which motion parallax depth cues are reliably obtained. It is known that similar data preprocessing is performed for visuomotor control tasks, and two dynamic mapping versions are advocated by various groups; either mapping into eye-centered coordinates, or into head-, and body coordinates. While for remembered target locations each of these mappings has its specific advantages, we show here, that optical flow is only simplified when dynamically mapped into head coordinates. There is little if any benefit for a depth-from-motion algorithm in a dynamic retinotopic map. Our results can be utilized in technical visual systems and we also suggest a verifiable hypothesis about a such a representation of optical flow in extrastriate cortex.

1 Introduction

One of the chief problems in computational vision is the three-dimensional reconstruction of a static scene from two-dimensional images [1]. Motion parallax is one of the depth cues that can be used to recover the three-dimensional structure of a viewed scene [2]. Motion induces a velocity field on the retina called the optical flow [3]. In the most general motion case, i. e., ego- plus object motion, the resulting curved optical flow field pattern cannot be resolved for depth analysis without additional assumptions [4] and even if simplifying assumptions are made, the problem of depth-from-motion remains rather complex.

Purely translational ego-motion induces one of the simplest optical flow fields. The optical flow has a fixed point, called the focus of expansion (FOE). All optical flow trajectories move outwards from the FOE. A radial flow field (RFF) contains reliable and rather easily accessible information about the three-dimensional structure of the viewed scene [5]. It is readily seen that the motion in such an RFF is one-dimensional in any retinotopic map—in a specific curvelinear coordinate system. For example, in retinal coordinates the RFF is expanding solely along the radial coordinate when an observer approaches an object. In coordinates of primate striate cortex this radial flow is mapped roughly along

parallel aligned neurons starting from the posterior pole to more anterior location in the medial occipital lobe [6]. We show in this study, that a flow field, with the only rotational components due to eye-gaze movements, is in a dynamic map isometric to a one-dimensional RFF. In other words, this flow filed is invariant to these specific rotational components. There are areas in extrastriate cortex known to have similar features as the dynamic map we suggest.

One of us (FW) introduced earlier an algorithm that efficiently analyzes an RFF, i.e., purely translational flow, and then reconstructs the viewed three-dimensional scene [5]. Details of the algorithm should be taken from that reference. We will use this algorithm to explore the use of a dynamic map—as described in the next section— for ego-motions with eye-gaze movements combined with straight body motion. We would like to emphasize, that any other depth-from-motion algorithms, that takes as input an one-dimensional RFF, can utilize the dynamical mapping strategy. However, the RFF-algorithm has been specifically designed to allow for parallelization of computations, and it is foremost this feature which is conserved by dynamic mapping.

2 Dynamical Mapping

To map the retinal flow field to a head centric frame, the retina is sampled by point-like receptive fields (Fig. 1, top layer). Initially, the receptive fields are placed such that they sample an RFF where direction of gaze and heading direction coincide. The receptive fields are positioned on a polar grid (receptive field grid, RFG) defined by radial axes expanding from the FOE. If the distance between successive receptive fields increases hyperbolically on each radial line, the optical flow is sampled uniformly.

The layout of the receptive fields on the RFG matches the radial optical flow field only if motion direction and direction of gaze coincide. When both directions differ by a constant angle α the receptive field positions on the RFG are re-mapped. After a gaze shift α about the Y-axis (angle of yaw), the optical flow is transformed by:

$$\theta^{hc}(\alpha) = f\sqrt{\left(\frac{\theta\cos(\phi)\cos\alpha - f\sin\alpha}{f\cos\alpha + \theta\cos(\phi)\sin\alpha}\right)^2 + \left(\frac{\theta\sin(\phi)}{f\cos\alpha + \theta\cos(\phi)\sin\alpha}\right)^2} \quad (1)$$

$$\phi^{hc}(\alpha) = \arctan\left(\frac{\theta\sin(\phi)}{\theta\cos(\phi)\cos\alpha - f\sin\alpha}\right)$$

The index hc indicates that these coordinates are head-centric while without index they are retinotopic.

In a head-centric frame the rotational component of the optical flow is counteracted by constantly updating the receptive field positions along with the rotational component according to the mapping function (Eq. 1). This is possible because direction and magnitude of the rotational component depend only on the angular velocity of the gaze change and not on any external information of the viewed scene. The optical flow in a head-centric frame is then congruent to

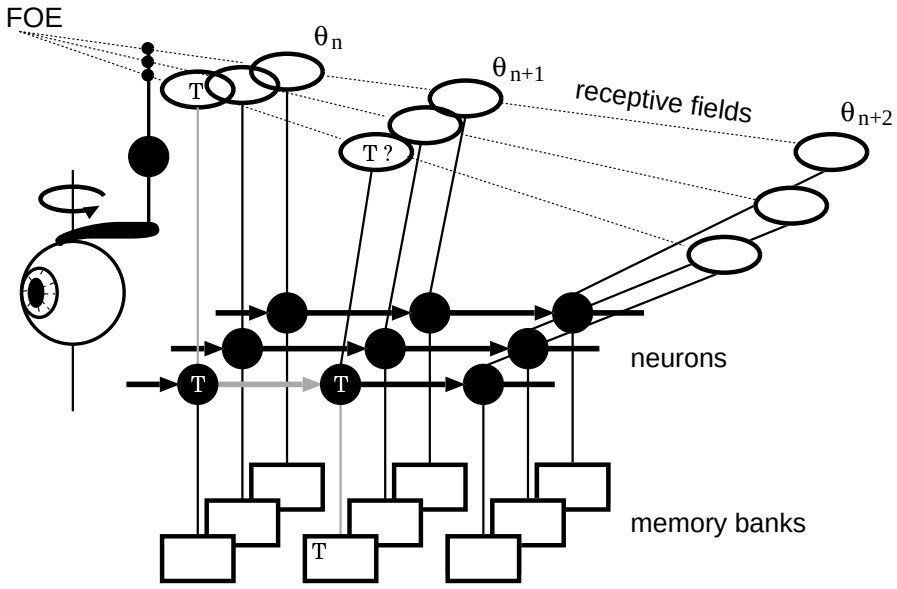

Fig. 1. Architecture of the three layer network. The top layer consists of receptive fields sampling the optical flow. Each receptive field projects to a neuron in the middle layer. The third layer consists of memory banks, one for each processing neuron. A separate neuron represents a structure mapping eye-positions. A visual tokens (T) is passed from the receptive field along the exemplarily shown grey connections towards the memory bank of a consecutive neuron. A head-centric representation of visual input in the middle neuronal layer is achieved by dynamically mapping the receptive field positions according to the direction of gaze. To re-construct three-dimensional position of viewed objects, the middle layer needs only locally exchanged information in one spatial direction (from left to right).

an RFF obtained with stable direction of gaze. Consequently, this dynamic map is invariant under eye-gaze movements and the RFF-algorithm can be applied on this head-centric map.

3 Performance of the RFF-algorithm on a head-centric map

An observer moving straight without changing the direction of gaze can adequately detect the three-dimensional position of the edges of objects in view by the RFF-algorithm [5]. For example, determining the distance of a teapot by the RFF-algorithm, results in three-dimensional coordinates, shown in front view (Fig. 2 A) and top view (Fig. 2 B). These detected coordinates outline the contour of the teapot. The depth coordinate Z, as shown in the top view (Fig. 2 B), is the actual output of the RFF-algorithm. The contour in the other

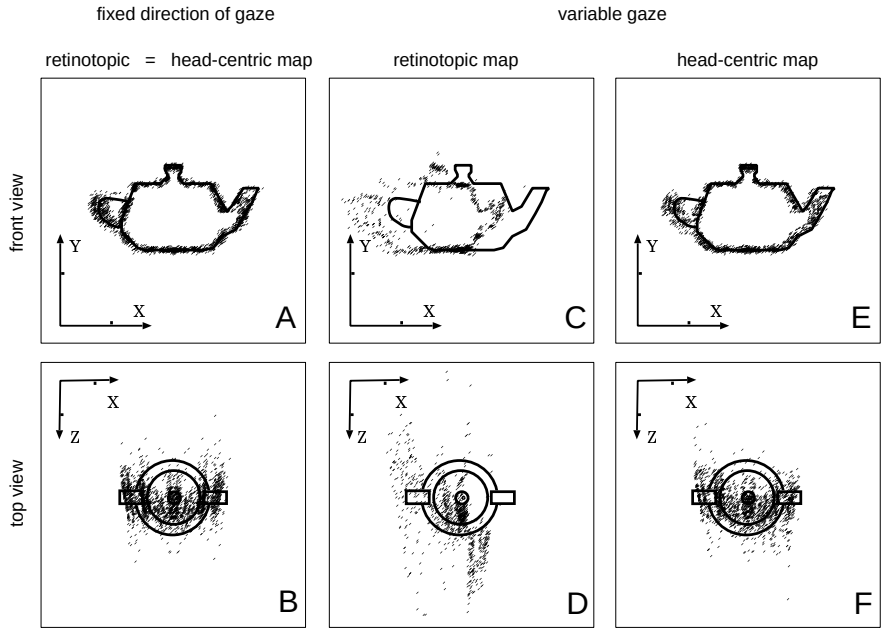

Fig. 2. A teapot viewed with stable and variable gaze. The position of the teapot in Cartesian coordinates (X,Y,Z) can be detected on a retinotopic map by the RFF-algorithm only when the gaze is pointing toward a fixed direction (A and B). Otherwise this algorithm makes systematic errors (C and D). If the position of the teapot is to remain stable, this algorithm must operate on a head-centric map (E and F). See also text.

two coordinates, X and Y (Fig. 2 A), are directly projected onto the retina and therefore they are already implicitly known, except for a scaling constant.

The detection of the teapot deteriorates when the straight body motion is combined with eye-gaze movements (front view Fig. 2 C, and top view D). There is even a shift of the projection of the teapot in the X-direction, that is, in the direction of one implicitly known coordinate (Fig. 2 C). This shift is inherent in the retino-centric map. Such a map can not statically store spatial locations. To be precise, edges of the teapot that are located on the retina right (left) from the FOE are accelerated (slowed down) by the additional rotational flow component, when the gaze rotates clock-wise about the Y-axis. This systematic change in the flow velocity is falsely interpreted by the RFF-algorithm as an edge too near (far), as shown by the tilt in Fig. 2 D. If the RFF-algorithm operates on head-centric optical flow fields, the performance of the RFF-algorithm is invariant under gaze sifts. (Fig. 2 E and F).

To quantify the performance of the RFF-algorithm on both the retinal flow field and the head-centric flow field, we defined a standard detection task: the three-dimensional reconstruction of a centric viewed square plane. For fixed direction of gaze along heading direction this corresponds to a situation where

edges move with hyperbolically increasing velocity along the receptive fields on each radial line. The angles between the edge and the radius vary between 0° and 45°. The average error in the detected three-dimensional position of the edges of the square plane was normalized to 1 for fixed direction of gaze (Fig. 3). If the gaze direction rotates stepwise by a total angle between 1° and 4° about the Y-axis, the error increases when the RFF-algorithm operates on retinal optical flow fields, as expected (see Fig 3). On head-centric optical flow fields the performance of the standard detection task is stable.

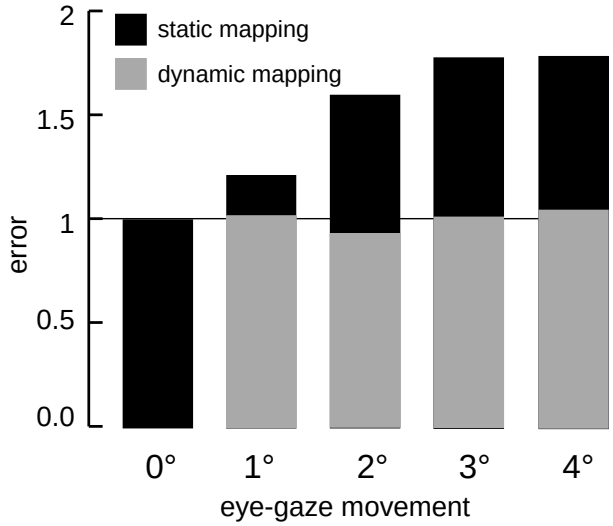

Fig. 3. Performance of the RFF-algorithm operating on a retinotopic map compared to a head-centric map. While on a head-centric map the performance is stable, on the retinotopic map it fastly deteriorates.

4 Discussion

Rotational components, foremost in form of smooth pursuit eye movements, are likely to occur in the ego-motion even within short periods of time. As soon as a rotational component is mixed with translation motion, the optical flow is two-dimensional in any coordinate system of a retinotopic map and extracting depth from optical flow becomes generally far more complicated. We showed that with a simple dynamic mapping strategy of visual space, the effect of eye-gaze movements on the optical flow can be eliminated. The resulting flow field on a head-centric map is congruent to the one induced by pure translational motion.

Dynamical mapping provides an example of combining two visual brain maps into one. In this case a subcortical sensor map that controls gaze direction in

retinal coordinates [7] and a retinotopic cortical map. The resulting map has qualitative new and advantageous features. We also attach importance to a head-centric map because it serves multiple though related purposes. In several areas in the parietal cortex: V3a [8], V5 [9], MST [10], V6 [11], V6a [12], 7a [13], and VIP [14] the activity of neurons is influenced by gaze direction. Precise gaze tuning together with a topographic representation of space can form a head-centric map. Area MST [15] and 7a [16] are both known to represent optical flow, although in different ways, and are likely candidates to utilize one-dimensional flow fields as depth cues, as we suggest here. To test this hypothesis, one needs to present radial expanding optical flow and introduce rotational components by pursuit gaze movements.

References

1. Marr, D. (1982). Vision. New York: W. H. Freeman and Company.
2. Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity-sensitive neurons, and space perception: a hypothesis. *Perception*, 3, 63–80.
3. Gibson, J. J. (1950). The perception of the visual world. Boston: Houghton Mifflin.
4. Poggio G.F., Torre V., Koch C.: Computational vision and regularization theory. Nature, **317** (1985) 314–319.
5. Wörgötter, F., Cozzi, A., & Gerdes V. (1999). A parallel noise-robust algorithm to recover depth information from radial flow fields. *Neural Computation*, 11, 381–416.
6. Tootell, R. B., Silverman, M. S., Switkes, E., & De Valois, R. L. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science*, 218, 902–904.
7. Klier, E.M., Wang, H., & Crawford, J. D. (2001). The superior colliculus encodes gaze commands in retinal coordinates. *Nature Neuroscience*, 6, 4 627-632.
8. Galletti, C., & Battaglini P. P. (1989). Gaze-dependent visual neurons in area V3A of monkey prestriate cortex. *Journal of Neuroscience*, 9, 1112–1125.
9. Puce, A., Allison, T., Bentin, S., Gore, & J. C., McCarthy G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18, 2188–2199.
10. Squatrito, S, & Maioli, M. G. (1996). Gaze field properties of eye position neurones in areas MST and 7a of the macaque monkey. *Visual Neuroscience*, 13, 385–398.
11. Galletti, C., Battaglini, P. P., & Fattori P. (1995). Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey. *European Jornal of Neuroscience*, 7, 2486–2501.
12. Galletti, C., Fattori, P., Battaglini, P. P., Shipp, S, & Zeki, S. (1996). Functional demarcation of a border between areas V6 and V6A in the superior parietal gyrus of the macaque monkey. *European Jornal of Neuroscience* 8, 30–52.
13. Andersen, R. A., Essick G. K., & Siegel, R. M. (1985). Encoding of spatial location by posterior parietal neurons. *Science*, 230, 456–458.
14. Duhamel, J. R., Bremmer, F., BenHamed, S., & Graf, W. (1997). Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature*, 389, 845–848.
15. Duffy, C. J., & Wurtz, R. H. (1995). Response of monkey M. ST neurons to optic flow stimuli with shifted centers of motion. *Journal of Neuroscience*, 15, 5192–5208.
16. Siegel, R. M., Read, H. L. (1997). Analysis of optic flow in the monkey parietal area 7a. *Cerebral Cortex* 7, 327–346.