Proc. Workshop "Dynamische Perception", Ulm, 2000, pp. 191-194, Akademische Verlagsgesellschaft / infix 2000

# Improvement of Optical Flow Estimates by Visuomotor Anticipation

**Volker Stephan, Horst-Michael Gross**

Department of Neuroinformatics, Ilmenau Technical University
98684 Ilmenau, Germany, P.O.B. 100565

**Abstract**

In this paper, we present a biologically inspired methodology to improve noisy estimates of optical flow by means of an internally generated expectation of the sensory consequenes of the systems own actions. The facilities of this anticipative preprocessing are demonstrated by means of an optical flow field based local navigation behavior of a miniature robot. Our anticipative preprocessing enables the robot to bridge gaps of sensory dropouts and, in consequence, to avoid collisions even with very noisy sensory information.

## 1 Introduction

In recent years, the appreciation of visual perception as a generative sensorimotor process gained increasing acceptance [3, 1]. The generative aspect of perception has been emphasized especially by [7, 8, 6] who supposed that internal simulation and mental imagery may play an integral role in perception. If this holds true at different levels of complexity and for different modalities, then, there must exist structures that are capable of predicting the sensory consequences of actions. Such sensory predictors seem to be multi-functional systems, since they can be used to a) enhance the incoming bottom-up sensory information by a top-down expectation generated previously b) direct selective attention to those environmental subregions which caused a mismatch of top-down expectation and bottom-up sensory information and c) internally simulate the consequences of action sequences in order to find and execute those actions, that entail positive outcomes for the system [5].
In this paper, we focus on a) and present a hybrid network architecture to enhance the incoming sensory information in a robot-navigation task.

## 2 Experimental framework

For our experiments, we use the real miniature robot platform KHEPERA equipped with an omni-directional color-camera (Fig. 1, left) to demonstrate the improvement on the incoming sensory information stream by fusion with an action-specific top-down expectation.
We believe, that the behavior of an autonomous system is a very good indicator of the performance

 

**Fig.** 1: *Used robot platform* KHEPERA *equipped with an omni-directional camera (left). Right: polar transformed image obtained from the omni-camera: middle=front, left and right image borders=back.*

of the system's 'perception' of its environment. The system's goal is a collision-free local navigation based only on visual information, in this case, the optical flow field. We use optical flow, because it is largely independent of specific visual details of the objects in the scene and yields implicit information about spatial distances to objects.
In the preprocessing of the original omni-camera-images we perform a polar transformation to the deskewed form depicted in Figure 1 (right). These transformed images are used directly to estimate the optical flow fields, because an action of the robot with a rotational part yields a rotation of the omni-camera-image but only a shift in x-direction of the polar transformed image. This is very advantageous, since the applied correlation based optical flow estimation [2] needs not cope with rotated correlation areas.

# 3 Architecture

We use a hybrid architecture to predict the optical flow fields as a result of the previous optical flow field and the real or hypothetical action to be executed. To demonstrate the functionality in a Khepera-scenario, we fuse the sensory bottom-up estimate and the top-down expectation in order to reduce the noise and gain robustness against sensory dropouts (see Figure 2).
A central aspect of our anticipative processing in the bottom-up/top-down cycle is the usage of flow vector specific confidence estimates organized topographically corresponding to the flow field. These confidence-values of each flow vector are based on correlation-based optical flow estimation [2] by evaluating the shape of the correlation function.

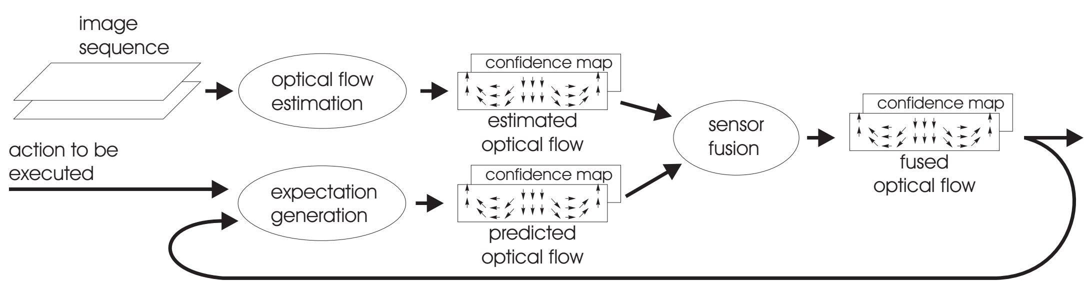


**Fig.** 2: *Hybrid architecture to fuse the sensory bottom-up and the top-down expectation.*

## 3.1 Expectation generation

Sensorymotor prediction is of central importance for our approach. In previous approaches [5], we used standard neural networks, such as multilayer-perceptrons. But, these globally operating networks had prediction problems due to the prohibitively high dimensionality of the sensory input, the whole optical flow field.
Hence, we developed an alternative approach to overcome this problem. It uses the property inherent to optical flow to represent the movements of objects within the camera-plane. Consequently, the inverse optical flow itself points to that part of the image, where movement must be predicted (see Figure 3). Thereafter, the current predicted vector results from the superposition of the source-vector and its previous value, both scaled in x and y-direction by weights, that are adapted during a learning period.

## 3.2 Fusion

With regard to Fig. 2, in this section we present the fusion between bottom-up and top-down information based on a small topologically coded 2D neural field, which is able to represent several

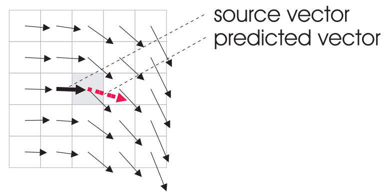


**Fig.** 3: *The flow vector to be predicted (dashed vector) depends only on those flow vector (bold), that point closest to the position for which the flow vector is to be predicted. That is, because this vector describes the velocity of the corresponding objects in the scene.*

alternative flow-hypotheses in parallel (Fig. 4). Consequently, both the sensory bottom-up and the top-down expectation can add their hypotheses about the real optical flow vector into the corresponding neural field, whereby similar hypotheses result in a superposition of the blobs at the same position. Hence, our fusion-algorithm utilizes a max-search to select those of all hypotheses, which support each other. This is reasonable, since similar information in both streams implies, that this information is reliable and trustworthy.
For reasons of simulation resources, we split the 2-dimensional neural field into 2 one-dimensional neural vectors representing the x- and y-direction of the flow vector separately.

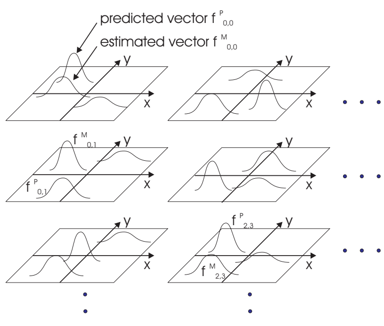


**Fig.** 4: *Each vector of the optical flow field is represented by a topologically coded 2D neural field, where the position within the neural field codes possible flow-vectors as blobs, and the activation of the blob is a measure for the corresponding confidence of that respective optical flow vector.*

# 4 Results

The flow field predictor was pretrained during a 3 minute random walk of the KHEPERA in its environment. To demonstrate the facilities of the presented anticipatory preprocessing, we placed the robot afterwards in unknown environments to navigate through a narrow passage without collision. For this benchmark, we used the balancing approach [4], which tries to equalize the optical flow in both hemispheres of the robot, which results in a collision-free locomotion (Fig. 5). If a perturbation is applied in this experimental situation, the usage of pure estimated optical flow fields fails, because the very noisy sensory input entails no information about close obstacles. In contrast, our anticipatory preprocessing allows the system to bridge the time gap of sensory dropouts with the generated action-specific expectation and is therefore able to extract relevant information in order to avoid the arising obstacles.

# 5 Conclusions and Outlook

In this paper, we presented a hybrid neural architecture to predict optical flow fields as consequences of own actions. Further, we introduced a neural field-based method to code and fuse sensory bottom-up estimates and top-down expectations.
The facilities of the anticipative preprocessing could be demonstrated by means of a local navigation behavior of the mobile robot platform KHEPERA. Future work will address the application of the sensory prediction for further tasks, such as to draw visual attention to moving objects, while the

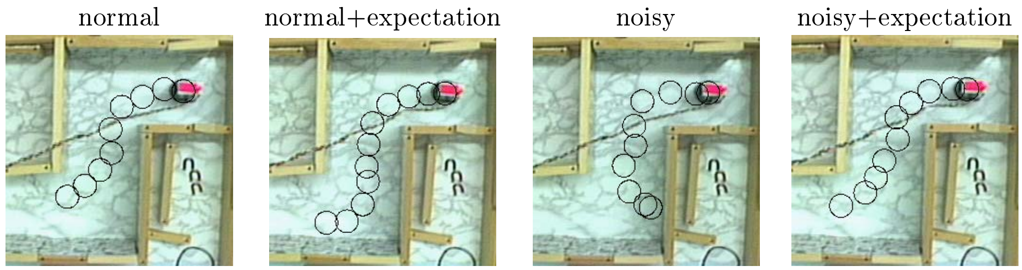


**Fig.** 5: *Navigation based on the estimated optical flow starting at the upper right corner and moving to the opposite one of the images. As can be seen, both the navigation on the pure estimated optical flow (first) and on the expectation driven preprocessed optical flow (second) allow a collision-free locomotion of the robot. In contrast, a significant disturbance of the optical flow estimate by fluctuating ambient light causes a collision at the end of the plotted trace, where no anticipative preprocessing is applied (third). The anticipative preprocessing overcomes the problems and allows a collision-free locomotion also with very noisy sensory information. (fourth).*

observer itself is moving, and evaluation of longer action sequences in order to find an optimal action sequence according to the current system state and goal [5].

Even though the presented architecture shows various similarities to the well known Kalman filter approach [9], there exist some essential differences about the fusion of a measurement and the corresponding prediction. In the case of the Kalman filter, the fusion is a weighted superposition of the two datastreams, whereas our architecture selects the most reliable hypothesis from the neural field. Thus, interpolation only occurs between sufficiently similar hypotheses, whereas very different hypotheses are not merged as they would be with the Kalman filter. Instead, the most trustworthy supposition is preferred. In consequence, our architecture is more robust against outliers. A detailed comparison with the Kalman-filter approach is subject of future work.

## References


[1] M.A. Arbib, P. Erdi, and J. Szentagothai. *Neural Organization: Structure, Function and Dynamics.* MIT Press, 1998.

[2] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of Optical Flow Techniques. *International Journal of Computer Vision, 12:1*, pages 43–77, 1994.

[3] D.T. Cliff. *Computational Neurothology: A Provisional Manifesto.* University of Sussex, School of Cognitive and Computing Sciences, 1990.

[4] A.P. Duchon, W.H. Warren, and L.P. Kaelbing. Ecological Robotics: Controlling Behavior with Optical Flow. *Proceedings of the 17th Annual Cognitive Science Conference. J.D. Moore and J.F. Lehman (eds.) Lawrence Erlbaum Associates.*, pages 164–169, 1995.

[5] H.-M. Gross, A. Heinze, T. Seiler, and V. Stephan. Generative Character of Perception: A Neural Architecture for Sensorimotor Anticipation. *Neural Networks*, 12:1101–1129, 1999.

[6] S.M. Kosslyn. *Image and Brain: The Resolution of the Imagery Debate.* MIT Press, 1996.

[7] S.M. Kosslyn, N.M. Alpert, and W.L Thompson. Visual mental imagery activates topographically organized visual cortex: PET investigations. *Journal of Cognitive Neuroscience*, 5(3):263–87, 1993.

[8] S.M. Kosslyn and A.L. Sussman. Roles of Imagery in Perception: Or, There is No Such Thing as Immaculate Perception. In M.S. Gazzangia, editor, *The Cognitive Neuroscience*, pages 1035–1042. MIT Press, 1995.

[9] G. Welch and G. Bishop. An Introduction to the Kalman Filter. Technical report, Department of Computer Science, University of North Carolina, 1997.