Unsupervised Learning of Sensory Primitives from Optical Flow Fields

Oswald Berthold* and Verena V. Hafner

Cognitive Robotics Group, Dept. of Computer Science
Humboldt-Universität zu Berlin
Rudower Chaussee 25, 12489 Berlin, Germany
{oberthol,hafner}@informatik.hu-berlin.de
http://koro.informatik.hu-berlin.de

Abstract. Adaptive behaviour of animats largely depends on the processing of their sensory information. In this paper, we examine the estimation of robot egomotion from visual input by unsupervised online learning. The input is a sparse optical flow field constructed from discrete motion detectors. The global flow field properties depend on the robot motion, the spatial distribution of motion detectors with respect to the robot body and the visual environment. We show how online linear Principal Component Analysis can be applied to this problem to enable a robot to continuously adapt to a changing environment.

Keywords: adaptive behaviour, source separation, feature learning, neural network, optical flow, primitives, redundancy, representation learning, sensor array, unsupervised, vision.

1 Introduction

Moving around in the world is *the* prime ability agents need for accomplishing things in a physical world. Many organisms have evolved to use some form of vision for sensing the motion of their bodies with respect to the environment and in relation to their own motor signals. The reason this works so well for animals also holds for robots. Their vision is fast, lightweight, passive and reliable through a large amount of redundancy.

Our approach to adaptive robot control is defined by learning data-driven primitives from raw sensorimotor channels [1]. These can be used for synthesis of behaviour in real world scenarios. We are motivated in this approach by the likely presence of similar organizational principles in biological nervous systems [2], [3].

The sensory information considered here is vision. A vision sensor is, at a fundamental level, an array, not necessarily homogenous, of photosensitive elements. This is true for both biological and technical systems. Motion is reflected in such an array as the propagation of a stable structure yielding spatio-temporally correlated excitation of the single elements. The role of motion detection in the visual sense, its implementation in early vision and the neural mechanisms underlying motion detection have been studied extensively in the literature [4]. Inspired by these ideas, many algorithms and circuits

* Oswald Berthold is funded by the DFG Research Training Group METRIK.

A.P. del Pobil et al. (Eds.): SAB 2014, LNAI 8575, pp. 188–197, 2014.

have been proposed that are able to locally detect elementary visual motion. An array of Elementary Motion Detectors (EMDs) comprises an Optical Flow (OF) field. While elementary motion is always local and planar, full egomotion can be estimated through wide-field integration. This is the reconstruction of the full 6 Degree of Freedom (DoF) motion parameters of the vision sensor (traveling with the animal or robot) in $\mathbb{R}^3$.

The problem can be solved through analysis of geometric properties of the sensor and environmental statistics. While the first part of this approach is straightforward [5], the latter part is not so easy to deal with because of large environmental variabilities. Another approach could be the computational modeling of known biological egomotion circuits and the implantation of these models [6] into an appropriate sensorimotor system (robot). We propose to follow an unsupervised learning approach on optic flow fields to solve this problem which make use of the high redundancy in the visual input to extract the underlying regularities which are imprinted on the raw sensory stream by camera geometry, viewing direction, motion type and environment all at once.

The paper is structured as follows: In section 2 we review related work. We go into more detail about the methods used for motion detection, signal acquisition, basis field extraction and recombination in section 3. We pick one method suitable for online learning and apply it in section 4 in simulation and on a real robotic car equipped with an onboard camera. We briefly discuss the results in section 5 and conclude with a summary in section section 6.

The main contribution of this paper is a demonstration of general unsupervised online learning of optical flow subspaces and respective component estimation on both real and simulated monocular cameras on different robotic vehicles.

2 Related Work

There are several research directions that contribute with respect to this problem, neurobiology, biorobotics, computer vision and machine learning. We will roughly group our quick survey of existing work according to these categories.

2.1 Biology and Biorobotics

Above we have mentioned neurophysiological work on decoding parts of the vision apparatus of insects leading to the concept of Elementary Motion Detectors (EMD). A summary on the topic is given in [4]. Here we are interested in properties of the global flow field, the total combined output of all EMDs in a vision system. In [7] the question of the principal resolvability of arbitrary 6 DoF motion through the visual input alone is brought up and answered affirmatively. The existence of specialized channels (primitives) in biological vision for basic orthogonal flow field components is hypothesized.

The question of ambiguity in motion fields is considered in [8] and found to be non-critical for practical concerns. Other work later expanded on these results [9] and compared the ideal motion field and the optical flow field. Their main argument is, that while quantitative equivalence can hardly be accomplished, its is only qualitative similarity which is of practical interest. The qualities there refer to the attractors of planar dynamical systems, which are used to model the basis flow fields.