Robotics and Autonomous Systems 71 (2015) 69–82

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Cortically inspired sensor fusion network for mobile robot egomotion estimation

Cristian Axenie *, Jörg Conradt

Neuroscientific System Theory Group, Department of Electrical Engineering and Information Technology, Technical University of Munich, Arcisstrasse 21, 80333 Munich, Germany

HIGHLIGHTS

- A cortically inspired sensor fusion network for robotic applications is introduced.
- Distributed graphical network, with combined feed-forward and recurrent connectivity.
- No global supervisor, only local processing, storage and exchange of information.
- Given sensory data, network relaxes into the best explanation of an underlying cause.
- Extensible to learn sensor correlations and adapt connectivity from incoming data.

ARTICLE INFO

Article history:
Available online 31 December 2014

Keywords:
Sensor fusion
Cortically inspired network
Egomotion estimation
Mobile robots

ABSTRACT

All physical systems must reliably extract information from their noisy and partially observable environment and build an internal representation of space to orient their behaviour. Precise egomotion estimation is important to keep external (i.e. environmental information) and internal (i.e. proprioception) cues coherent. The constructed representation subsequently defines the space of possible actions. Due to the multimodal nature of incoming streams of sensory information, egomotion estimation is a challenging sensor fusion problem. In this paper we present a distributed cortically inspired processing scheme for sensor fusion, which given various sensory inputs, and simple relations defining inter-sensory dependencies, relaxes into a solution which provides a plausible interpretation of the perceived environment. The proposed model has been implemented for egomotion estimation on an autonomous mobile robot. We demonstrate that the model provides a precise estimate of both robot position and orientation.

1. Introduction

An essential component in motor planning and navigation, for both real and artificial organisms, is egomotion estimation. Egomotion or self-motion refers to the combined rotational and translational displacement of a perceiver with respect to the environment. During motion organisms build their spatial knowledge and behaviours by continuously refining their internal belief about the environment and own state [1–3]. Our approach is motivated by three main aspects consistent with recent results in spatial processing for navigation and perception [4].

The first aspect addresses the importance of maintaining a precise position of the self. Building an internal representation of the environment and own state implies the coherent alignment of the acquired sensory cues. As sensory cues are conveyed from both dynamic egomotion related signals such as odometry and inertial signals, and static external environmental signals, such as visual or auditory, the precise position of the self links and keeps the representation coherent. In this context a coherent representation provides the ability to recognise and define "action possibilities" from all available sensory cues (e.g. distance to objects). Subsequently, egomotion defines the space of possible actions and impacts behaviour [2,3,5].

A second aspect refers to the capability of a real or artificial organism to understand space itself from its own state (in space). Egomotion estimation contributes to the understanding of high-level features of the environment, like structure and layout, such that the organism can direct actions and control its movement. Typically, with respect to position, the primary question is related to distances to key objects in the environment. In order to infer correct distances, the organism must traverse the environment and

---

* Corresponding author. Tel.: +49 89 289 26908.
E-mail addresses: cristian.axenie@tum.com (C. Axenie), conradt@tum.de (J. Conradt).
URL: http://www.nst.ei.tum.de (C. Axenie).

http://dx.doi.org/10.1016/j.robot.2014.11.019
0921-8890/

distinguish between its dynamic and static features as they lead to different consequences [6].

The third aspect points directly to the solution offered in this paper, namely how can precise egomotion perception be obtained given the complex multisensory environment. In order to handle environmental variability and complexity, continuous and simultaneous incoming sensory data streams from different sensors must be combined into a robust representation. But sensory cues are usually complementary and redundant and it is not clear how they describe the spatio-temporal properties of the environment. To disambiguate the complex scenario the global representation should combine all cues in an informative and plausible way. This combination process, termed sensor fusion, is not trivial, as current implementations [7–9] show. The primary objective is aligning reference systems of the different congruent and redundant sensory cues. After alignment, depending on inferred spatio-temporal correlations, interference and conflicts between the cues need to be minimised [2,3,10]. Finally, sensor fusion should not propagate biases or errors in the final (fused) estimate but compensate for them.

The paper is organised as follows. After a review of existing sensor fusion mechanisms and the motivation for the bioinspired paradigm shift in Section 2, we introduce our model in Section 3. Starting from the general neurally inspired processing model, we present the architecture and the specific instantiation for the mobile robot egomotion estimation. Section 4 provides the analysis and evaluation of our model and a comparison with state-of-the-art methods, whereas Section 5 provides a thorough discussion of the obtained experimental results. Finally, Section 6 concludes the work, by summarising the main features and advantages of the model and introducing future extensions.

2. Review of sensor fusion algorithms for egomotion estimation

Currently developed engineered approaches for sensor fusion typically aim at optimal solutions, and many results in robotic applications demonstrate this [7–9]. Real-world scenarios, however, are typically dynamic, prone to parameters changes and characterised by complex features. To cope with such aspects, typical approaches often need parameter tuning or model refinements. Their dedicated structure cannot handle the variability of the percepts or accommodate different scenarios. These algorithms cannot easily handle different contexts from those considered in the design. When considering adaptivity and robustness, neurobiology offers vastly superior performance over today's engineered systems. With different processing paradigms and distributed representations our brain solves the task of combining sensory information not only more effectively, but seemingly without much effort. Extremely flexible, the brain can easily accommodate and handle, by learning, new and different tasks and situations.

Having set up the framework, in the upcoming sections we first provide an overview of state-of-the art sensor fusion approaches and their instantiations using specific computational methods. Second, we mark the advantages of the paradigm shift towards bioinspired computational approaches. Supported by examples in robotic cognitive systems we introduce the main features of neural systems which transferred to technical systems will provide a higher degree of flexibility and robustness.

2.1. *Standard computational methods for sensor fusion*

Most state-of-the-art sensor fusion algorithms are based on probabilistic models of observations and processes. This framework has become attractive for both engineering and computational neuroscience as a powerful tool to describe sensory models and dynamics (in engineering) [7–9] and also inference and sensory integration in populations of neurons (in computational neuroscience) [11–15].

These algorithms use Bayes' rule to integrate observations and system's model into a unified estimate of the system's state. In addition, these methods replace point representations of perceptual estimates with probability distributions such that the statistics of the sensory estimates can be quantified and used for inference. Bayesian methods provide an optimal estimation scheme, in the sense that their estimate of the given state is unbiased (i.e. difference between this estimator's expected value and the true value of the state being estimated is zero/estimation is true on average) and has minimum variance [7–9]. In Bayesian inference belief is progressively updated as new data from the sensors is presented such that the initial belief evolves towards an informed posterior distribution. In many simple cases it is possible to analytically describe the posterior distribution. This is the case where the prior and likelihood function are given by Gaussian normal distributions [7–9,16]. However, in many real-world scenarios this is not the case [16,15].

As an alternative approach to probabilistic approaches other non-Bayesian approaches have been developed. Providing a strong framework to describe uncertainty by using the notion of partial membership, fuzzy logic, is a powerful tool for imprecise reasoning [8]. Although fuzzy theory is particularly useful to represent and fuse information provided by human experts, it is limited merely to fusion of vague data [17]. More often this method is integrated with probabilistic fusion algorithms [18]. Using a qualitatively different reasoning technique,the Dempster–Shafer theory of evidence, fuses information relying on probability mass to characterise belief and plausibilities in the data. Despite the ability to fuse uncertain and ambiguous data, evidential reasoning is inefficient in fusing highly conflicting data and has scaling problems in the case of high-dimensional state spaces [8,19]. Even though a large variety of methods to fuse sensory data were developed, they are usually combined to improve performance in more general scenarios.

Independent of the underlying computational method in use, current sensor fusion algorithms' implementations are described by a global and sequential processing scheme. Dictated by current computing architectures this paradigm constrains algorithms to obey to a pipelined sequence of filters and other feed-forward processing stages. Neuroscience studies postulate that neural processing is described by distributed and unsupervised computation mechanisms, with mixed feed-forward and recurrent flow of information and local storage and processing capabilities [11–14]. These mechanisms can adapt to novelty, by learning, and exhibit robustness in the face of uncertainty. Transferring these principles to technical systems will support this paradigm shift towards more flexible and adaptive systems.

2.2. *Bioinspired approach to sensor fusion: changing paradigm*

Sensor fusion is a process that influences major aspects of perception, cognition and behaviour in both physical and artificial systems [1,7,8,20]. Traditionally sensor fusion describes a mechanism to combine cues from the same or different modalities, converted to a common, internal representation which is subsequently used in the actual fusion process [21,22]. It is commonly agreed that the brain contains areas specialised for processing different types of information incoming from sensors [23–25]. A major determinant for a brain area's ability to process a certain type of information is the input it receives [23,26]. It is considered that the unique processing characteristic of each cortical area is defined in terms of the area's interactions with the other areas [27,23,26,12–14]. Hypotheses from cortical interareal coordination studies support evidence that these areas aim to reach a consensus and maintain mutually consistent information with the others resolving coherence or incoherence relations (i.e. constraints) [26,28,12]. Furthermore, neurobiological evidence supports the view that elementary