Robotics and Autonomous Systems 71 (2015) 1–2

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

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

Editorial

Emerging spatial competences: From machine perception to sensorimotor intelligence

1. Introduction

Following the recent evolution of robotics and AI in different fields of application, the increasing complexity of the actions that an artificial agent needs to perform is directly dependent on the complexity of the sensory information that it can acquire, represent and interpret, along the perception process.

From this point of view, an efficient internal representation of the sensory information is a grounding factor of a robot to develop a human-like capability of interaction with the surrounding environment. Moreover, in the space at a reachable distance, not only visual and auditory, but also tactile and proprioceptive information rise to be relevant to gain a comprehensive spatial cognition. This information, coming from different senses, can be in principle integrated and used to experience an awareness of the environment both to actively interact with it, and to calibrate the interaction itself. Besides, the early sensory and sensorimotor mechanisms, that at a first glance may appear to be simple algorithms, are grounded on highly structured and complex processes that are far from being understood and modeled. By exploiting an early synergy between sensing modules and motor control, the loop between action and perception comes to be not just closed at system level, but also shortened at an inner one. This would allow not only the emergence of spatial competences, but also their continuous adaptation to changes in the environment or in the body, which could modify its interactions with the world.

2. Focus of the special issue

The aim of this special issue is to survey a state of the art of concepts, methodologies, techniques and algorithms that would serve as bricks on which to build and develop artificial agents with spatial competences, where perceptual and cognitive understanding of space emerge from sensorimotor exercise. It stems from a special session that were organized within the 2013 International Joint Conference on Neural Networks (IJCNN), held in Dallas (TX), USA, 4th–9th of August 2013. The purpose of the special session, named *From Sensing Machines to Sensorimotor Intelligence* was to challenge the Scientific Community on the entwined nature of the sensory representation and the sensorimotor control of an active robot agent. More details about the event can be found on the official INNS website at: http://www.inns.org/ijcnn-past-conferences.

Following the interest in the topic received at IJCNN, we promoted this Special Issue, which was open not only to papers presented at the Special Session but also to other contributions not previously published. This special issue collects the 13 best papers selected among the 32 papers submitted, *i.e.* with an acceptance rate of 40%. We believe that they are interesting examples of intelligent behavior that directly emerges from the interaction with the environment.

3. Contributions

This special issue presents a diverse range of works that can be grouped in three main areas, ranging from learning sensorimotor skills for visual active exploration of space, through action-minded space representation, up to the interpretation of space through actions. It includes contributions that present novel approaches, as well as papers that illustrate the topic to a broad, interdisciplinary audience.

3.1. Sensorimotor skills for visual active exploration of space

A branch of studies focuses on active vision systems, and in particular on different kinds of eye movements, *i.e.* smooth pursuit, saccade and vergence. These works show how eye movements must be continuously recalibrated in order to maintain their accuracy, and both the learning and calibration processes not only strictly rely on, but also contribute to form the representation of the visual information. **Teuliére et al.** rely on a model for the smooth pursuit of moving objects, where the motor commands for tracking and the sensory representation are learned in parallel. The algorithm combines an efficient coding approach for learning sensory representations with reinforcement learning scheme, which seeks to minimize the encoding reconstruction error between two consecutive frames. Using this approach, the system learns smooth pursuit of moving objects without the need for a calibration procedure, and surprisingly, it achieves a good performance both on simulation and on real robot. **Antonelli et al.** investigate different approaches to learn an internal model for saccade control in humanoid robots, and discuss the performance in terms of accuracy, adaptation and space representation. Gathering inspiration from the functionality of the cerebellum and its interaction with the brain-stem, the resulting model, tested on an anthropomorphic robot, performs an implicit representation of space, given visual and somato-sensory information. **Gibaldi et al.** rely on a cortical architecture, inspired by the complex cells of the primary visual area

http://dx.doi.org/10.1016/j.robot.2015.03.006
0921-8890/

of primates, to control vergence eye movements. Interestingly, the proposed learning rule relies on a reward signal representing the variation of the overall activity of the population, which allows for a continuous and autonomous learning without any external supervision.

3.2. Action-minded internal representation, and negotiation of space

A proper control of eye movements is at the base of the visual perception, and accordingly of the capability of the sensing agent to perform a correct interaction with it. Once this *early* competence is developed, a higher *intermediate* interpretation of the scene is required to decide which action to perform. **Ji et al.** introduce a brain-inspired developmental architecture that simulates the phenomenon of attention and recognition, grounded on the interaction of a bottom-up and a top-down learning mechanism. The implemented approach is able to learn a bidirectional relationship between the position and type of an object falling on the retinal image. Aiming at a system that is able to *learn* not just the visual exploration and interpretation of the scene, but also a physical interaction with it, a direct *experience* of the surrounding environment is a necessary step to gain a comprehensive spatial cognition. Indeed, the arm or the body itself, being part of the sensing agent, can be used as exploration instrument to gain a spatial sensorimotor map, where the body and its sensory representation both act as a reference metric of the space. **Terekhov et al.** propose a method that allows an agent to build an internal representation of the location of its end effector, independently of the configuration of the intermediate effectors, providing the agent with an intrinsic notion of space. Indeed, the sensor experience is exploited not only to start creating a structure of the external space, but also its internal representation. **Conradt and Stewart** implement a spiking neural network on a multi-sensor mobile robotic platform. The information provided by the different sensors (vision, bump, virtual movement) during the exploration of an unknown environment is fused to obtain a sensorimotor map that allows for an effective and real-time navigation capability. **Axenie et al.** present a distributed cortical inspired processing scheme for sensor fusion, providing a plausible interpretation of the perceived environment. The obtained interpretation is thus used to obtain a precise estimate of a robot egomotion, in terms of position and orientation. **Droniou et al.** describe a deep self-organized system that is able to extract sensory motor contingencies from the various sensory modalities, like proprioception, vision and audition, assessed on a humanoid robot. The implemented hierarchical architecture allows for bidirectional prediction among any number of modalities and enables the robot to learn how to write and read written symbols through human kinesthetic teaching.

3.3. Interpretation of space through actions

Once the sensing agent can rely on effective capabilities for exploring the surrounding space and to bridge the sensor modality with the motor action, *complex* competences are necessary for a coherent interaction in space. A meaningful approach to undertake is to interpret the space from a functional perspective, with respect to the possible actions that are or can be performed. **Mojtahedzadeh et al.** tackle with the difficult and highly relevant problem of determining support relations among quasi-randomly stacked objects for safe manipulation. The authors solve the problem for known environments with a geometric reasoning method estimating the stationary equilibrium of objects, while for unknown environments they use a statistical classifier to infer which object supports which other one. **Aksoy et al.** address the problem of on-line, incremental learning of the semantics of manipulation actions observed from human manipulation. A semantic event chain is used to compress the information regarding the sequence of changes of the spatial relations between objects, allowing the learning of complex manipulation actions. **Lee et al.** describe a human intent recognition system using the combination of two supervised recurrent neural networks on multiple time scales. The lower layer is used for action recognition whereas the second layer is used for intent recognition on the top of the first one. **Hawng et al.** deal with the topic of mimicking human motion by three-dimensional information extracted via a stereo vision system. Once a skeleton capture algorithm identifies the joints position in space, a hybrid learning model is used on the basis of inverse kinematics of the body segments to classify and imitate the human motion. **Khasnobish et al.** present an integrated approach to bridge robotic sensation and human perception, allowing a human user to feel through a robot arm. An artificial arm is able to recognize embossed digits by means of a pressure distribution image acquired by exploring the embossed area, and the digit is transmitted to the user who operates the artificial arm through vibrotactile stimuli.

Acknowledgments

The guest editors wish to express their gratitude to the Program Committee members of the Special Session as well as all the additional reviewers who provided helpful assistance in the overall revision process with valuable suggestions and comments. We also would like to thank the Editor in Chief of Robotics and Autonomous Systems, Prof. Rüdiger Dillman, for the opportunity to publish this Special Issue.

Agostino Gibaldi *
Silvio P. Sabatini
Physical Structure of Perception and Computation Group, Department of Informatics, Bioengineering, Robotics and System Engineering, University of Genoa, Italy
E-mail addresses: agostino.gibaldi@unige.it (A. Gibaldi), silvio.sabatini@unige.it (S.P. Sabatini).

Sylvain Argentieri
Institute for Intelligent Systems and Robotics (ISIR), Sorbonne Universités, UPMC University, Paris 06, France
E-mail address: sylvain.argentieri@upmc.fr.

Zhengping Ji
Advanced Image Research Laboratory (AIRL), Samsung, Pasadena, CA, USA
E-mail address: jizhengp@gmail.com.

* Corresponding editor.