The modular modality frame model: continuous body state estimation and plausibility-weighted information fusion

Stephan Ehrenfeld · Martin V. Butz

Received: 9 March 2012 / Accepted: 25 September 2012 / Published online: 23 October 2012
© Springer-Verlag Berlin Heidelberg 2012

Abstract Human show admirable capabilities in movement planning and execution. They can perform complex tasks in various contexts, using the available sensory information very effectively. Body models and continuous body state estimations appear necessary to realize such capabilities. We introduce the Modular Modality Frame (MMF) model, which maintains a highly distributed, modularized body model continuously updating, modularized probabilistic body state estimations over time. Modularization is realized with respect to modality frames, that is, sensory modalities in particular frames of reference and with respect to particular body parts. We evaluate MMF performance on a simulated, nine degree of freedom arm in 3D space. The results show that MMF is able to maintain accurate body state estimations despite high sensor and motor noise. Moreover, by comparing the sensory information available in different modality frames, MMF can identify faulty sensory measurements on the fly. In the near future, applications to lightweight robot control should be pursued. Moreover, MMF may be enhanced with neural encodings by introducing neural population codes and learning techniques. Finally, more dexterous goal-directed behavior should be realized by exploiting the available redundant state representations.

Keywords Modularity · Multimodality · Plausibility estimation · Body schema · State estimation · Robot arm control

1 Introduction

Although humans face many challenges when interacting with the environment, they master them with great skill. This dexterous control capability (Latash and Turvey 1996) can in part be accounted for solely by the body and hardwired reactive control architectures. Examples are embodied robots presented by Braitenberg, Brooks, or Schmitz, which can produce astonishingly versatile behavior (Braitenberg 1986; Brooks 1990; Schmitz et al. 2008). However, these embodied systems are limited to stereotypic behavior. They cannot generate the full dexterity humans demonstrate.

Instead, a body model is necessary for the realization of highly flexible behavior. The notion that humans develop various internal body models is commonly accepted. Various researchers have suggested that these body models are based on forward and inverse representations of sensorimotor correlations and also include estimations of the own body metric (Bernier et al. 2007; Denève et al. 2007; Hoffmann et al. 2010; Longo and Haggard 2010; Streri et al. 1993; Wolpert and Kawato 1998). Humans appear to selectively exploit these representations to predict, plan, and control movements. Thus, suitably modularized bodily representations are crucial for the dexterity apparent in human motor planning and control (Bernier et al. 2007; Latash and Turvey 1996; Latash et al. 2007; Rosenbaum 2010; Wolpert and Kawato 1998).

Nikolai A. Bernstein pointed out that the dexterous control capability of humans must be due to a highly flexible and modular neural representation and control system, which effectively deals with many redundant degrees of freedom of the human body (Latash and Turvey 1996). A key to human dexterity thus lies in the architecture of neurally encoded body models. In the following, we further elaborate on this conjecture. In particular, we propose that a suitable body model
should represent its internal state estimations probabilistically, it should be capable of flexibly integrating multisensory information, and it should maintain redundant state estimations in multiple body-part and sensory-motor respective modules.

1.1 Probabilistic state estimations

Uncertainty arises not only during motor execution and due to sensor noise, but also due to uncertainty about the internal body model, the body state, the task, and the state of the environment. Examples are unpredictable changes in the environment caused by other agents and physical influences. Also the body schema, bodily kinematics, and dynamics change during growth (Wells et al. 2002), injury, bodily exercise, or food intake (Shadmehr and Wise 2005). All these sources of uncertainty give rise to a complicated challenge, as the representations of body, goals, and constraints depend—either directly or indirectly—on the sensory system, which cannot generate noiseless, absolutely precise information. Humans address this uncertainty by means of probabilistic body state estimations (Körding and Wolpert 2004; Doya et al. 2007).

1.2 Multisensory integration

Humans gather state information with a wide variety of sensors in different modalities (like vision and proprioception) and integrate it in their perception of their own body (e.g., Beauchamp 2005; Makin et al. 2008; Maravita et al. 2003). Thus, redundant sources of information interact in the human brain in various ways (Calvert et al. 2004; Tononi et al. 1998). These interactions can lead to peculiar effects, such as the rubber hand illusion (Botvinick et al. 1998; Makin et al. 2008), where a representation in one modality frame is strongly modified by information stemming from other modalities and frames of reference. Sensory information from different modalities is thus flexibly combined selectively and probabilistically, dependent on current accuracy and plausibility estimates.

1.3 Modularized representations

Seeing that different sources of information are available in different brain areas, body and environment are selectively and redundantly represented in many state estimation modules. In addition, the brain modularizes its state estimations body-part respectively. For example, bodily self-perception is clearly separated into body segments (de Vignemont et al. 2009). Also, not only primary motor cortex and somatosensory cortex show bodily homunculi, but also many parietal cortical areas and premotor areas exhibit body part distinctions and selective, multimodal, sensory, and reafferent information source integration (Andersen et al. 1997; Shadmehr and Krakauer 2008). With respect to behavior, different parts of the body and bodily components can be moved by selectively activating motor synergies (Gentner and Classen 2006; Latash et al. 2007). With respect to particular movement tasks, a modularized redundant representation allows immediate selective access to the most relevant information amongst the redundant alternatives. Thus, highly modularized, redundant state estimations of body and environment are relevant both, for a robust and flexible representation of state estimations and for dexterous behavioral decision making and control.

Inspired by these characteristics, we propose the Modular Modality Frame (MMF) model. MMF is a highly modularized architecture, which maintains a probabilistic body state estimation over time in redundant modules. Forward and inverse kinematic mappings link these modules and allow for a flexible flow of information, resulting in sensor fusion and local interaction of multiple state estimations. The benefits of our model include high noise robustness, low dimensional representations, and the ability to estimate sensory reliability on the fly.

In Sect. 2, we describe the general architecture of the proposed model: its modality frames and the cooperation of different steps in maintaining the body state representation. These steps are described in more detail in Sect. 3. They include movement prediction, sensor fusion using plausibilities, sensor integration, and crosstalk between modules. The calculation of sensor plausibilities is defined in Sect. 4. The necessary local forward and inverse kinematic mappings are detailed in Sect. 5. In Sect. 6 the experimental setup is described, and results are presented and discussed. The work is concluded with a short summary and a discussion of interesting potential system extensions and applications.

2 MMF model overview

MMF is inspired by the knowledge about the architecture of the human brain and human behavioral control capabilities, as sketched-out above. In particular, MMF represents an internal body model with modularized, locally interactive representations. Each module maintains a particular, probabilistic body state estimation over time by integrating the available multisensory information by means of Bayesian principles.

The body model comprises a full forward kinematics model (used for prediction) and a full inverse kinematics model (used for planning and control), providing high accuracy in simple goal-directed movements. For extensive overviews over other body models in robotics cf. Hoffmann et al. (2010), Nguyen-Tuong and Peters (2011), Sigaud et al. (2011). MMF’s modules differ from each other with respect to (I) the arm limb, (II) the sensory modality, which can be a perceived location, direction, or joint angle, and (III) the frame of reference, which can be “global” (i.e., head-