New Ideas in Psychology 31 (2013) 221–238

Contents lists available at SciVerse ScienceDirect

New Ideas in Psychology

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

Robot studies on saccade-triggered visual prediction

Wolfram Schenck

Computer Engineering Group, Faculty of Technology, Bielefeld University, POB 10 01 31, 33501 Bielefeld, Germany

Keywords:
Visual prediction
Motor learning
Eye–arm coordination
Mental imagery
Cognitive modeling
Robotics

A B S T R A C T

Three robot studies on visual prediction are presented. In all of them, a visual forward model is used, which predicts the visual consequences of saccade-like camera movements. This forward model works by remapping visual information between the pre- and post-saccadic retinal images; at an abstract modeling level, this process is closely related to neurons whose visual receptive fields shift in anticipation of saccades. In the robot studies, predictive remapping is used (1) in the context of saccade adaptation, to reidentify target objects after saccades are carried out; (2) for a model of grasping, in which both fixated and non-fixated target objects are processed by the same foveal mechanism; and (3) in a computational architecture for mental imagery, which generates “gripper appearances” internally without real sensory inflow. The robotic experiments and their underlying computational models are discussed with regard to predictive remapping in the brain, transsaccadic memory, and attention. The results confirm that visual prediction is a mechanism that has to be considered in the design of artificial cognitive agents and the modeling of information processing in the human visual system.

1. Introduction

The prediction of future states is an important aspect of cognition. Without prediction, coordinated behavior would be impossible, at any level of cognition. Take, as one extreme, social cognition: To interact in a meaningful way with other people, human actors have to predict other peoples’ reactions to their own actions (Wolpert, Doya, & Kawato, 2003). At the other extreme, in sensorimotor processing, prediction is required in the nervous system to generate ongoing internal feedback during dynamic movements during which the feedback from proprioception is too slow (Miall & Wolpert, 1996). Generally, prediction encompasses the generation of information about future states of the body and the environment. There is a large body of evidence, from neuroscience and psychology, that the brain is continuously engaging in such predictive processing (Bubic, von Cramon, & Schubotz, 2010).

Visual prediction is of special interest for many biological and artificial agents because of the important role vision plays in navigation, obstacle avoidance, object recognition, social interaction, etc. On the one hand, there is a need to predict how observed external events will unfold in the future; for instance, a predator has to anticipate the movements of its prey, maybe even of single limbs, from previously gathered visual data. On the other hand, prediction serves to anticipate the consequences of one’s own motor actions; that is, how the visual sensory inflow will change as a result of eye, head, or whole body movements. In what follows, only the second type of prediction will be considered.

The focus of this paper is on visual prediction triggered by saccadic eye movements. It is written from the perspectives of robotics and computational modeling. The starting point is a phenomenon called “predictive remapping” (Wurtz, 2008). This term refers to neurons in various brain regions that have one special characteristic: Their visual receptive field shifts *before* saccade onset from its default location to a different location in the field of view. This is exactly at the spot which will be projected onto their default receptive field *after* the saccade is executed. This remapping of receptive fields can be interpreted as

E-mail address: wschenck@ti.uni-bielefeld.de.

0732-118X/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.newideapsych.2012.12.003

a predictive process. Interestingly, remapping has also proved to be a robust and feasible solution for visual prediction in *technical* systems. Schenck and Möller (2007) developed a visual forward model[1] for a robot camera head. The model predicts how the visual sensory inflow will change as a consequence of camera movements. Since the model was also intended to be biological in scope, "retinal images" were employed which mimic the cone distribution on the retina. This study showed how the faithful prediction of raw visual data can be achieved, and, further, how an adaptive visual forward model can be trained for this task. The two main problems of visual prediction were consequently solved: first, that a large number of pixels or sensors leads to high-dimensional input and output of visual predictors, and second, that part of the future visual state is not predictable at all because it is out of sight at the time of prediction. For these reasons, only a few studies in the area of robotics have dealt directly with visual prediction, most of them with mobile robots (Gross, Heinze, Seiler, & Stephan, 1999; Hoffmann, 2007; Hoffmann & Möller, 2004).

Starting from predictive remapping and the existing visual forward model, three robot studies are sketched in this paper. Each study employs saccade-related visual prediction in a different way for a different motor, perceptual, or cognitive task. The underlying computational models are not directly related to psychological studies and do not strive to explain specific experimental data sets, but they put forward hypotheses as to how information processing could be carried out by the brain in some domains. These hypotheses undergo their first test in real-world robot experiments; after passing this test, they are intended as new ideas for further psychological research. Thus, this paper is not about the technical mechanisms of predictive remapping, which have been described before (Schenck, Hoffmann, & Möller, 2011; Schenck & Möller, 2007), but on how visual prediction based on remapping can be used for different tasks.

The paper is organized as follows. In Section 2, neurophysiological and psychological studies on predictive remapping are briefly reviewed. In Section 3, the robot setup and the visual forward model are described. Afterward, three different applications of visual prediction are presented. The first application explores how the target reidentification problem in saccade learning can be overcome by visual prediction (Section 4). The second presents an architecture for the grasping of extrafoveal targets, which relies on visual prediction triggered by shifts in spatial attention (Section 5). The third addresses the problem of mental imagery (Section 6).

2. Predictive remapping

"Predictive remapping," as we previously noted, refers to neurons which respond in anticipation of a saccade to a stimulus which will appear in their visual receptive field (RF) only after the saccade has taken place. This process is depicted in Fig. 1. This response characteristic is usually interpreted as a temporary shift in the retinal location of the RF. After the saccade, the RF switches back to its default location. In his comprehensive review, Wurtz (2008) lists more than ten studies in which neurons with such predictive responses were found in monkeys. They were identified in the lateral intraparietal area (LIP; e.g., Duhamel, Colby, & Goldberg, 1992; Heiser & Colby, 2006), in the frontal eye fields (FEF; e.g., Umeno & Goldberg, 1997), in earlier extrastriate visual areas (e.g., Nakamura & Colby, 2002), and in the superior colliculus (SC; e.g., Walker, Fitzgibbon, & Goldberg, 1995).

In addition, remapping of the response to visual stimuli has been investigated in human beings. Merriam, Genovese, and Colby (2007) carried out a fMRI study that tracked activity in the visual cortex during a single-step saccade task. They found activity which could not be attributed to the visual stimulus or to oculomotor processing alone, but which resulted from the conjunction of these sensory and motor events and was interpreted as response to the remapped stimulus trace. In their data, the remapping response was stronger the higher the position in the visual hierarchy (more precisely, stronger in cortical regions hV4 and V3A than in regions V2 and V1). However, because of the temporal limitations of fMRI Merriam et al. were only able to conclude that remapping occurred around the time of the saccade, not whether it was a predictive response. To overcome this limitation, Parks and Corballis (2010) recorded EEG data during a single-step saccade task in which the hypothesized remapping would occur either cross- or intra-hemispherically (the distinction had to be coarse because of the low spatial resolution of the EEG method). The analysis of the event-related potentials revealed a presaccadic remapping response, which continued during saccade execution.

Evidence for predictive remapping stems also from studies at the behavioral level. Melcher (2007) conducted psychological experiments on how the tilt aftereffect (TAE) was modulated by a saccade between two fixation points. Participants were initially adapted to a tilted grating pattern overlying the first fixation point. After a short delay, the fixation point moved to its second position, triggering the saccade. The TAE was measured at times around saccade onset, either at the first or second fixation point. The TAE moved to the second fixation point even before saccade onset (60% increase at the second target position, 80% decrease at the first position). This finding is of special interest since it shows that a specific feature (not just the general activation level) was remapped in a predictive way.

For predictive remapping to occur, information about the new position of the eye has to be available. Because the eye has not moved yet, proprioception cannot provide this information. The only source is the scheduled but not yet executed saccadic motor command. A "copy" of this command (often termed corollary discharge [CD]; Sperry, 1950) has to drive the remapping process. Neurons whose activity correlates with saccade generation have been found in various brain areas, and the CD in primates is embedded in a multitude of neural pathways (Wurtz, 2008). It is not clear yet how the neurons with shifting RFs in different

[1] In the framework of internal models (Miall & Wolpert, 1996), a forward model predicts future sensory states in response to the current sensory state and a planned motor command.