Action Understanding and Imitation Learning in a Robot-Human Task*

Wolfram Erlhagen[1], Albert Mukovskiy[1], Estela Bicho[2], Giorgio Panin[3], Csaba Kiss[3], Alois Knoll[3], Hein van Schie[4], and Harold Bekkering[4]

[1] Dept. Mathematics for Science and Technology, University of Minho, 4800-058 Guimaraes, Portugal
{wolfram.erlhagen, am}@mct.uminho.pt
[2] Dept. Industrial Electronics, University of Minho, 4800-058 Guimaraes, Portugal
estela.bicho@dei.uminho.pt
[3] Informatics, Chair for Robotics and Embedded Systems, Technical University Munich, 85748 Garching, Germany
{panin, knoll, kiss}@in.tum.de
[4] Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, 6500 HE Nijmegen, The Netherlands
{H.vanschie, H.bekkering}@nici.ru.nl

Abstract. We report results of an interdisciplinary project which aims at endowing a real robot system with the capacity for learning by goal-directed imitation. The control architecture is biologically inspired as it reflects recent experimental findings in action observation/execution studies. We test its functionality in variations of an imitation paradigm in which the artefact has to reproduce the observed or inferred end state of a grasping-placing sequence displayed by a human model.

1 Introduction

In robotics research imitation has attracted a lot of attention in recent years since it is considered a promising learning mechanism to transfer knowledge from an experienced teacher (e.g. a human) to an artificial agent. Most work has been focused on motor learning paradigms in which the imitating robot has to match as close as possible the kinematics and dynamics of an observed movement (for review see [1]). However, a growing body of experimental findings in imitation studies with humans indicate that the imitator most likely does not encode the full detail of the observed motions but the interpretation of those motions in terms of the demonstrator's goal. Very often differences in embodiment (e.g., child-adult) and/or task constraints (e.g., obstacles) simply do not allow for a matching on the level of the movement trajectory. In the goal-directed theory of imitation proposed by Bekkering and colleagues [2,3],

* The work was supported by the European grant *ArteSImit* (IST-2000-29686). We would like to thank Drs Leonardo Fogassi and Giacomo Rizzolatti for numerous discussions about this work.

W. Duch et al. (Eds.): ICANN 2005, LNCS 3696, pp. 261–268, 2005.
© Springer-Verlag Berlin Heidelberg 2005

imitative behavior can be considered successful whenever the end state of the witnessed action is reproduced. The action means, on the other hand, may or not coincide with the observed ones.

In the work reported here: (a) we present a robot control architecture for goal-directed imitation which reflects processing principles discovered in recent experimental findings of action observation/execution studies, (b) we propose a biologically plausible learning scheme for establishing the links between means and goals during development and practice, (c) we test the complete control architecture in variations of an imitation paradigm in which a robot tries to reproduce the observed or inferred outcome of a grasping and placing sequence displayed by a human, (d) we show that knowledge about the meaning of an object may be transferred to the robot by imitation.

Goal-directed imitation requires of course that the imitator understands the action of the model. The neuro-cognitive mechanisms underlying action understanding are currently topic of an intense debate (e.g., [4,5]). A growing body of empirical evidence supports the notion that the production and perception as well as the interpretation of others' actions rely on a common distributed neural system. The 'direct matching hypothesis' proposed by Rizzolatti and colleagues [5] based on their discovery of the mirror system states that an action is understood when its observation activates the motor representations controlling the execution of the same action. However, the correspondence problem between dissimilar embodiments challenges an explanation purely based on a simple and direct resonance phenomenon of the motor system. Moreover, humans and monkeys are able to infer goals without a full vision of the action by integrating additional contextual cues.

The proposed model architecture for action understanding and goal-directed imitation in artefacts is based on the theoretical framework of dynamic fields [6,7]. It aims at implementing the idea that inferring motor intention is a continuous process which combines sensory evidence, prior task knowledge and a goal-directed matching of action observation and action execution.

2 Experimental Setup

For the robotics work we adopt a paradigm which has been developed to experimentally investigate in humans the idea that actions are organized in a goal-directed manner (van Schie and Bekkering, in preparation). The paradigm contains two objects of different color that must be grasped and then placed at one of two laterally presented targets that differ in height. The possible hand trajectories are constrained by an obstacle in form of a bridge (see Panel A in Fig. 1). Depending on the height of the bridge, the lower target may only be reached by grasping the object with a full grip and transporting it below the bridge. Placing the object at the higher target, on the other hand, may require combining a precision grip and a hand trajectory above the bridge.

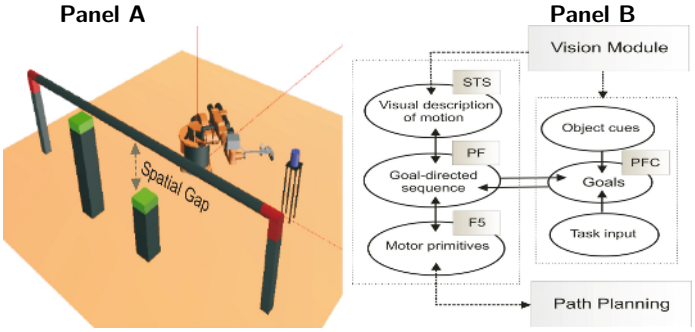

Fig. 1. Panel A: Bridge Paradigm. **Panel B:** Robot control architecture.

3 The Control Architecture

The artefact used in the imitation study consists of an industrial robot arm (Kuka with 6 DOFs) with a four-fingered anthropomorphic robot hand (Graal-Tech, University of Genova) and a real-time vision system. Three interconnected modules define the robot control architecture (Panel B in Fig. 1).

The *vision module* provides the environmental variables of the task setting (shape and position of bridge, position of object and targets etc.) by means of a semi-automatic calibrated stereo camera system. In addition, it tracks the detected object(s) and the hand of the instructor, classifies the demonstrated action in terms of grip and trajectory type, and identifies the placing target. All outputs are globally available for the other modules.

In the *cognitive module*, decisions about the action goal and the means to achieve that goal are made. Its layered architecture is biologically inspired, as it represents the basic functionality of neuronal populations in interconnected brain areas known to be involved in action observation/execution tasks (for details see [8]). The core part consists of three reciprocally connected layers, STS, PF and F5, representing the mirror circuit. The fundamental idea is that within this circuit the matching of action observation and action execution takes place on the level of motor primitives which abstract from the fine details of the movements [5]. Concretely for the bridge paradigm, we distinguish two types of grasping primitives (precision (PG) and full (FG) grip) and two types of transporting primitives for avoiding the obstacle (below (BT) or above (AT) the bridge). The visual description of the observed action is stored in STS. In the motor layer F5 the representations of the respective primitives become active both during action observation and action execution, that is, we assume that those primitives already exist in the motor repertoire of the robot. The representations in the intermediate layer PF reflect recent neurophysiological findings in brain areas PF/PFG that suggest a goal-directed organization of action means. Using a sequence task, Fogassi and colleagues [9] described a population of grasping neurons which showed a selective response in dependence of the final goal of the action (eating or placing) to which the grasping act belongs. For the bridge

paradigm, we abstract this finding by assuming representations of specific combinations of primitives (e.g., PG-AT) which allow achieving a specific goal. Layer PF is reciprocally connected with a prefrontal area (PFC) in which the goals parameterized by their height relative to the bridge are encoded. A goal representation may be triggered or influenced by visual input (placed object), through the links to PF, and/or learned associations to representations of object cues (e.g., color) and memorized task information (e.g., number and probability of goals).

In the *path planning module*, the abstract primitives of layer F5 are translated into a movement plan generating the right kinematics. We assume that the path planning takes place in posture space. This requires that a model of the inverse kinematics for the arm-hand system is known. For the planning we employ the framework of wave expansion networks [10] with nodes representing stored arrays of joint angles. The sequence of postures defining a collision-free path for the robot arm-hand is found by propagating an activity wavefront between nodes encoding the initial and the desired goal postures. Posture nodes which are impossible due to the obstacle are inhibited. They are found by explicitly testing for spatial overlap in Cartesian space between the to-be-assumed posture and the bridge (forward maps). Moreover, the ensemble of nodes which can become part of the wavefront is further constrained by the motor primitives in F5. For instance, we use again forward maps to check whether a particular node represents 'all links in a high position' as required by an AT-trajectory. This integration of prior information together with the inherent parallelism of the wavefront operations makes a real-time path planning for artefacts with higher degrees of freedom feasible.

4 Hebbian Learning of the Synaptic Links

Each layer of the cognitive module is formalized by a Dynamic Field [6,7] in which self-sustained patterns of excitation encode task specific information. The layer dynamics is governed by the following equation:

$$\tau \frac{\delta}{\delta t} u(x,t) = -u(x,t) + g(u(x,t)) \left[\int w(x-x') f(u(x',t)) dx' - \right.$$
$$\left. -w_{inhib} \int f(u(x',t) dx' \right] + h + \sum_i S_i(x,t) \qquad (1)$$

where $\tau > 0$, $h < 0$ and $w_{inhib} > 0$ are constants. The non-linear functions $f(u)$ and $g(u)$ are of sigmoid shape, the excitatory connections, $w(x, x')$, are modelled as a Gaussian profile. The excitation patterns evolve under the influence of multiple information sources, $\sum_i S_i(x,t)$, representing input from the visual module and from excitation in connected layers. Recurrent inhibition in each layer creates a competition between response alternatives (e.g., type of grasping) and guarantees the stability of the decision process represented by the excitation patterns.

To develop the goal-directed organization within the distributed network model, the synaptic links between the various layers have to be established during practice. We apply a correlation based learning rule for the synaptic connections, $a(x, y)$, between any two neurons x and y belonging to two different layers (for a discussion in the context of action understanding see [4]):

$$\tau_s \frac{\delta}{\delta t} a(x, y, t) = -a(x, y, t) + \eta \, f(\bar{u}_1(x)) f(\bar{u}_2(y)) \tag{2}$$

where $\tau_s \gg \tau$, $\eta > 0$ and $\bar{u}_1$, $\bar{u}_2$ denote the equilibrium solutions of the relaxation phase in layer 1 and layer 2, respectively. Importantly, we assume that an internally generated reinforcement signal representing a successful path planning toward the desired goal posture defines the time window for the learning. As a result, the matching of action observation and action execution becomes goal-directed, since the metric for the learning is not defined by the similarity in the kinematics but the similarity in the end state [11].

5 Experimental Results

A set of imitation experiments within the Bridge paradigm has been performed which differ mainly in the amount and nature of the visual information available to the robot. In the first experiment, a complete visual description of the teacher's action in terms of the grasping and transporting behavior exists and

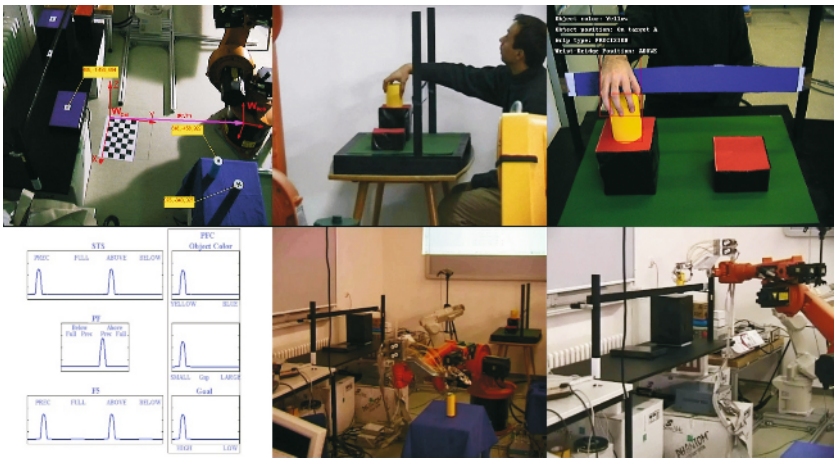

Fig. 2. Example of an imitation task in which the robot reproduces the placing on the higher goal using the means displayed by the human teacher (PG-grip, AT-trajectory). The upper row depicts the visual analysis of the world configuration (left) and the teacher's action (right). On bottom, decisions for goal and means represented in the dynamic field model (left) and two snapshots of the robot in action are shown.

the visual system identifies the goal. The visual description in STS resonates via the matching mechanism in the mirror circuit with the congruent motor primitives in PF and F5. If the covert path planning toward the desired goal-posture turns out to be successful, the observed response strategy can be associated with the goal for future use by the learning procedure described above. Figure 2 illustrates the result of this learning by imitation in an example in which the robot copies the demonstrated precision grip and the trajectory above the bridge to place the object at the higher goal. By using objects with different properties (e.g., color), the robot may acquire additional knowledge in repeated imitation trials by learning an association between object cues and the goal where to place a particular object ('object meaning'). For instance, yellow objects have to be placed at the higher and blue objects at the lower target.

The second experiment shows that the learned link from the mirror circuit to the goal representation is crucial. The bar of the bridge is removed for the human teacher but not for the robot (Panel A in Fig. 3). Because of this change in the environmental constraints, the teacher now uses a full grip for placing the yellow object at the higher target. For the robot, a direct matching on the level of motor primitives would result in a collision with the bridge. As shown in the snapshot of the stable state of the network dynamics in Panel A of Figure 3, the decisions in layer F5 represent the motor primitives PG (grip) and AT (trajectory) previously associated with the higher goal parameterized by the smaller

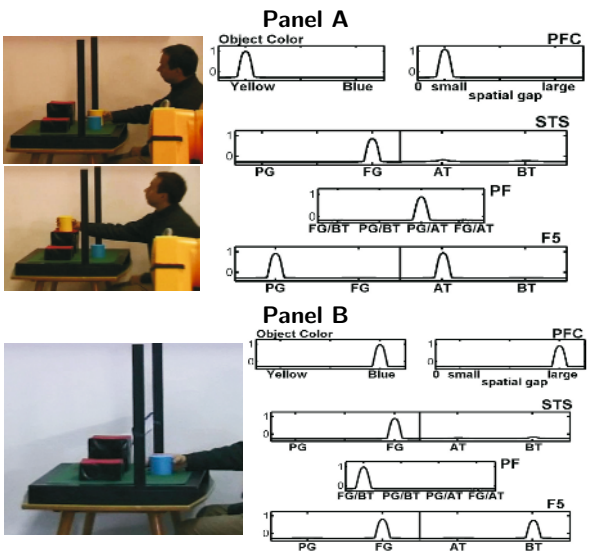

Fig. 3. **Panel A:** Conflict in the grasping behavior: the teacher uses a full grip for placing the object. As shown in layer F5 of the field model, the robot decides to use a precision grip to reproduce the end-state. **Panel B:** Inference task: only the grasping behavior is observable, the transpor ting and placing is hidden from view. The stable state in layer PFC of the field model represents the inferred goal.

spatial gap relative to the bridge. The observation of the placing triggers the respective goal representation which in turn biases the decisions in PF and F5 toward the PG-AT sequence. The input from PFC may override the direct mapping in the mirror circuit since in known task settings the goal representations in PFC evolve faster compared to the visual motion representations in STS (for details see [8]).

The third experiment reflects a situation which is very common for agents acting in cluttered and dynamic environments. Only partial visual information about the action displayed by another agent is available and the observer has to infer the action goal by integrating additional contextual cues and prior task information. Again we assume that the underlying mechanism for discerning motor intention is a goal-directed motor simulation. Consistently, it has been recently reported that grasping mirror neurons fire when the crucial final part of the demonstrated action is hidden but the monkey knew that there is a graspable object behind the occluding surface (for review [5]). In Panel B of Figure 3 we show a snapshot of the model dynamics in an inference task in which only the grasping with a full grip was observable and the rest of the action was hidden from view. Despite the missing visual information, the dynamics has relaxed in each model layer to a stable peak solution. The representation of the lower goal (parameterized by the larger spatial gap relative to the bridge) is triggered via the learned STS-PF-PFC pathway. Note that for the specific height of the bridge, the FG-BT sequence is the only sequence including the FG-grip which is associated with a particular goal. If the robot and the human model already share the knowledge about the meaning of the object, the color information serves as an additional, redundant input for the goal representation in PFC.

6 Discussion

The experiments with the real robot system illustrate some of the advantages of a goal-directed organization of imitative behavior compared to other current approaches which emphasize a matching on the trajectory or path level [1,11]. It allows coping with differences in embodiment and task constraints known as the correspondence problem [12] in robot imitation. Most importantly, it enables the robot to infer the purpose of the observed movement which is crucial for transferring specific knowledge from the model to the imitator. The idea that the movement production system is essentially involved in action understanding has been proposed in the context of robotics research several times in the past (for review [1]). For instance, Demiris and Hayes [13] used internal forward models known from motor control theory to predict the sensory consequences of observed actions in an imitation task. However, the use of forward models implicitly requires that imitator and demonstrator share similar bodies. Wolpert and colleagues [14] have recently proposed a computational scheme which also includes hierarchically higher levels of motor control in the simulation loop for action understanding. On this view, the goal-directed organization of action means in layer PF may be seen as part of an abstract forward model for

interpreting an ongoing action. We have recently shown that the goal-directed control architecture may even allow acquiring the meaning of an observed motor act which is not strictly in the motor repertoire of the observer [8].

In our current work, we consider paradigms that involve a richer set of motor primitives and more complex goal-directed sequences.

References

1. Schaal, S., Ijspeert, A., and Billard, A.: Computational approaches to motor learning by imitation. Phil. Trans. R. Soc. Lond. B Vol. 358 (2003)537-547
2. Bekkering, H., Wohlschläger, A., and Gattis, M.: Imitation of gestures in children is goal-directed. The Quartely Journal of Experimental Psychology, Vol. 53A (2000) 153-164
3. Wohlschäger, A., Gattis, M., and Bekkering, H.: Action generation and action perception in imitation: An instantiation of the ideomotor principle. Phil. Trans. R. Soc. Lond. B. Vol. 358 (2003)501-515
4. Keysers, C., and Perrett, D. I.: Demystifying social cognition: a Hebbian perspective. Trends in Cognitive Science. Vol.8 (2004)501-507
5. Rizzolatti, G., Fogassi, L., and Gallese, V.: Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews. Vol.2 (2001)661-670
6. Amari, S.: Dynamics of pattern formation in lateral-inhibitory type neural fields. Biological Cybernetics. Vol.27 (1977)77-87
7. Erlhagen, W., and Schöner, G.: Dynamic field theory of movement preparation. Psychological Review. Vol.109 (2002)545-572
8. Erlhagen, W., Mukovskiy, A., and Bicho, E.: A dynamic model for action understanding and goal-directed imitation. Submitted
9. Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. and Rizzolatti, G.: Parietal lobe: from action organization to intention understanding. Science. Vol.308 (2005) 662-667
10. Kassim, A., and Kumar, B.V.K.V.:The wave expansion neural network. Neurocomputing. Vol.16 (1997)237-258
11. Billard, A., Epars, Y.,Calinon, S., Schaal, S. and Cheng, G.: Discovering optimal imitation strategies. Robotics and Autonomous Systems. Vol.47 (2004)69-77
12. Alissandrakis, A., Nehaniv, C.L., and Dautenhahn, K.: Imitation with ALICE: Learning to imitate corresponding actions across dissimiliar embodiments. IEEE Transactions on Systems, Man, and Cybernetics- Part A. Systems and Humans. Vol.32 (2002)482-496
13. Demiris, J., and Hayes, G.: Imitation as a dual-route process featuring predictive and learning components: A biologically plausible computational model. Imitation in Animals and Artifacts. Dautenhahn, K. and Nehaniv, C.L. MIT Press, Cambridge MA (2002)327-361
14. Wolpert, D.M., Doya, K. and Kawato, M.: A unifying computational framework for motor control and social interaction. Phil. Trans. R. Soc. Lond. B. Vol.358 (2003)593-602