in Working Notes, Workshop on Agents in Interaction-Acquiring Competence Through Imitation, 2nd International Conference on Autonomous Agents (Agents '98), Minneapolis/St. Paul, May 10-13, 1998

Perceptuo-Motor Primitives in Imitation

John Demiris
Department of Artificial Intelligence
University of Edinburgh
Edinburgh, Scotland, UK
johnde@dai.ed.ac.uk
http://www.dai.ed.ac.uk/students/johnde/

Maja J Matarić
Computer Science Department
University of Southern California
Los Angeles, CA 90089-0781
mataric@cs.usc.edu
http://www-robotics.usc.edu/~maja/

Abstract

In this paper we examine the issues involved in selecting a set of perceptuo-motor primitives to be used in an imitation learning framework. We advocate a strong link between perception and action in imitation, and based on neuro-physiological and psychophysical data, we present a set of primitives for movement imitation. We describe our implementations of these primitives on a humanoid simulator with dynamics, and discuss how learning by imitation can take place in such a framework.

1 Introduction

Imitation is one of the most powerful forms of social learning, and one of the techniques frequently employed by human infants in order for them to benefit from the experience of their adult caretakers (Meltzoff, 1996). Given its frequent occurrence in nature, it is not surprising that roboticists (Hayes and Demiris, 1994, Kuniyoshi et al, 1994, Dautenhahn 1995, Demiris et al, 1997) have sought to utilize it in order to equip robots with the ability to imitate and learn from demonstration. From the cognitive modelling point of view, imitation is also interesting since it requires the interaction of several cognitive systems, including those of perception, memory, and motor control.

This paper discusses an approach to modeling imitation that emphasizes a strong link between the perceptual and motor systems. We describe background evidence from human and monkey neurophysiological and psychophysical data that supports our approach, and we present a set of movement primitives that can be used for both perception and production of movement. We have started implementing them on a realistic physics-based simulator of a human torso that includes dynamics. We describe how learning can take place in such a framework.

2 Perceptuo-motor interactions

Psychophysical experiments with human subjects have shown a close link between the perception and the production of an action. Viviani and Stucchi (1992) among others, have shown that actions are perceived under the strong influence of the motor capabilities of the observer. Even more persuasively, Rizzolatti and his colleagues (Rizzolatti et al, 1996) have found "mirror" neurons in the pre-motor cortex of monkeys, that are activated both by the perception and execution of actions. The majority of the neurons were selectively active during the observation and execution of a particular action only and did not respond to the sight of the object that an action might have been directed to, or other control demonstrations. These data indicate that, at some level, perceptual and motor systems share the same representational substrate, supporting and extending Positron Emission Tomography data (Decety et al, 1994) which indicated that passive observation of movements and mental imagery share common neural mechanisms with aspects of motor control, including planning.

3 A set of primitives

The idea behind movement primitives is that the complex and high-dimensional articulated control problem could be addressed by structuring the motor system as a collection of *primitives* which can then be sequenced and combined to produce the complete and complex general repertoire of movement. Biological support for such an organization of the motor system is provided by Bizzi et al (1991), Mussa-Ivaldi et al (1994) and their continuing work. This approach stands in contrast to the explicitly planning approach, which would compute a trajectory each time one is needed; with primitives, stereotypical trajectories are looked up (and possibly parameterized by the specific task), rather than computed *de novo*.

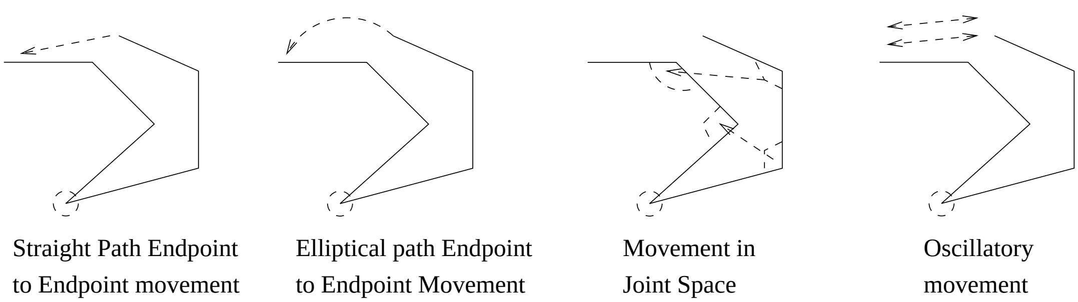

Figure 1: The selected set of movement primitives, executed by a 3-joint arm.

The issue of which movement primitives to choose is one of the most fundamental in this approach. We have selected a set of four, as shown in Figure 1, motivated by their different external characteristics, and by biological evidence indicating that they are likely handled in different ways by the motor system. The primitives are:

- *Straight-path EndPoint to EndPoint Movement:* This primitive implements a straight path movement of the endpoint of the controlled effector in Cartesian space. Movements of this form include reaching towards a target, and placing the endpoint into a specific position. This type of movements has known physical characteristics including bell-shaped velocity profiles (Flash and Hogan, 1985). There is also biological evidence (Matarić and Pomplun, 1997) that the human perceptual system, while observing movements of human demonstrators, focuses its attention to the end point of the moving limb. Note that this primitive deals with unobstructed paths. The next addresses collision and obstacle avoidance.
- *Elliptical-path EndPoint to EndPoint Movement:* This primitive generates an elliptical path of the endpoint of the controlled effector. Movements of this form include reaching to a target while avoiding an object. This type of movements is hypothesized to be planned through a selection (according to some optimization criterion, such as minimum-motor-command change, minimum torque-change, etc) of a set of via-points through which the effector's end-point will move, and then smoothed.
- *Posture-Achieving Movement in Joint space:* This primitive implements a movement that achieves a certain posture in joint space. Movements of this form are used in dance, gestural communication, etc.
- *Oscillatory Movement:* This primitive implements a periodic movement defined for either the end effector or the joints. Movements of this form include repetitive/oscillatory movements like walking, juggling, and bouncing. This type of movement is hypothesized to be executed by some type of a central pattern generator (CPG) system.

4 Implementation

We have started implementing the above set of primitives in a realistic rigid-body simulator of a human torso, with static graphical legs ("Adonis", Figure 2) which includes the dynamics of movement. Adonis has eight rigid links: head, torso, left and right upper arms, lower arms, and hands. The links are connected with rotary joints of three degrees-of-freedom (DOFs) in the neck, shoulders, and wrists, and one degree-of-freedom pin joints in the elbows. The waist is connected to the stationary legs through a three DOF joint. In total, Adonis has 20 DOFs. Mass and moment-of-inertia information is generated from the graphical body parts and equations of motion are calculated using a commercial solver, SD/Fast (SD/Fast User Manual).

So far, we have implemented the following primitives in Adonis:

- *Straight-Path EndPoint to EndPoint Movement:* Implemented (by Matthew Williamson) using impedance control (Hogan, 1985), and described in detail in Matarić et al (1998).
- *Posture-Achieving Movement in Joint space:* Implemented by comparing the current and target angle values of each joint, and calculating the forces that are needed to be applied at each joint to reach the desired joint angles.

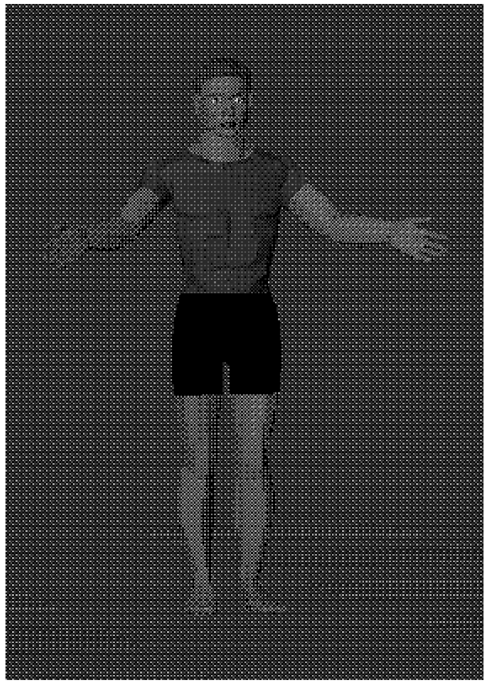

Figure 2: The humanoid dynamics simulator, Adonis

5 Discussion

How is the proposed primitive-based system used within an imitation learning framework? We envision the following approach: the agent is initially given a set of "innate" primitives. Whenever it observes other agents' movement, it maps the observed movements into its own set of primitives and internally (and externally, if needed) imitates them, by creating a movement plan consisting of a sequence of its own primitives. If the external characteristics of the reconstructed task match the observed ones, then imitation was successful. Otherwise, a process of learning occurs whereby the agent attempts various combinations of the primitives it has (including executing some of them in parallel), until the match between the characteristics of the generated and the observed processes is maximized.

We have started implementing this approach within a specific task: learning to dance the "Macarena" (Matarić et al, 1998). The Macarena is a dance that consists of a sequence of mostly upper body movements (a hip-movement and whole-body turn were omitted), which makes it an ideal task for Adonis. The sequence of movements includes extending the arms straight out, rotating the arms, placing the arms behind the head, etc. Interested readers can find a complete verbal specification of the Macarena at `http://www.radiopro.com/macarena.htm`. Importantly, the task includes both postural and end-point-control primitives, both straight and elliptical. We have been conducting experiments with human subjects in an attempt to determine whether features of the position and velocity profiles of the body parts can be used in order to recognize the different subparts of the task. Preliminary results indicate that this is possible by selecting as candidates the segments between points where the velocity of one or more body parts reaches zero.

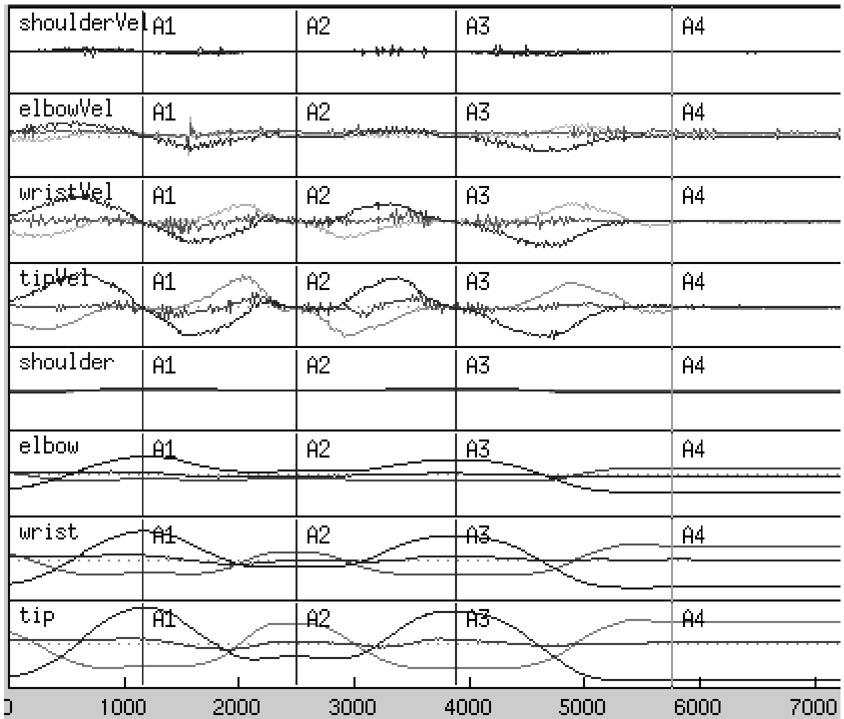

Figure 3: Position and velocities profiles for the example movement

For example, Figure 3 illustrates the position and velocity (signals postfixed with 'Vel') profiles of body points of a simple movement sequence:

To A1: Starting from the arm rest position next to the body, raise the right arm 90 degrees to the side keeping the arm straight during the movement (i.e., shoulder-elbow and elbow-wrist segment on a straight line).

To A2: Lower the elbow-wrist segment to reach 90 degrees with respect to the shoulder-elbow segment. This should bring it to a line parallel to the body.

To A3: Reverse the previous movement - bring elbow-wrist segment back to A1.

To A4: Bring the arm to the rest position, keeping the arm straight for the duration of the movement.

The human subject was asked to move at normal speed; the sequence took approximately seven seconds to complete. The data were captured using Qualisys' MacReflex motion analysis system, with uses five infrared cameras placed around the subject, and small reflective markers at the center of the chest, center of the waist, shoulder, elbow, wrist, and tip of the middle finger of the subject's right arm.

We have also obtained data of the positions and angles that are needed to reach the subgoals (postures) of the Macarena by programming Adonis to go through a preselected sequence of movements, and are currently attempting to recognize the subgoals, in order to implement the above-described imitation system.

Acknowledgments

The authors thank Stefan Schaal, Victor Zordan, Matt Williamson, Aswath Mohan, and Gillian Hayes for discussions on the issues addressed in this paper. Thanks also to Heba Lakany for her help in capturing the human movement data.

References

Emilio Bizzi, Ferdinando A. Mussa-Ivaldi, and Simon Giszter (1991), "Computations Underlying the Execution of Movement: A Biological Perspective", in *Science*, vol 253, pp. 287–291, 1991.

Kerstin Dautenhahn (1995), "Getting to Know Each Other - Artificial Social Intelligence for Autonomous Robots", in *Robotics and Autonomous Systems*, 16, pp. 333-356, 1995.

J. Decety, D. Perani, M. Jeannerod, V. Bettinardi, B. Tadary, R. Woods, J. C. Mazziota and F.Fazio, 1994, "Mapping Motor Representations with Positron Emission Tomography", in *Nature*, vol. 371, pp. 600–602, October 1994.

J. Demiris, S. Rougeaux, G. M. Hayes, L. Berthouze, Y. Kuniyoshi, 1997, "Deferred Imitation of Human Head Movements by an Active Stereo Vision Head", in *Proc. of the 1997 IEEE Workshop on Robot-Human Communication*, Sendai, Japan, Sept. 1997.

Tamar Flash and Neville Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model", in *Journal of Neuroscience*, 5(7):1688-1703, July 1985.

Gillian Hayes, and John Demiris, 1994, " A Robot Controller Using Learning by Imitation". in *Proc. of the 2nd International Symposium on Intelligent Robotic Systems*, Grenoble, France, July 1994.

Neville Hogan, 1985, "Impedance Control: An Approach to Manipulation", in *Journal of Dynamic Systems, Measurement, and Control*, 107: 1–24, 1985.

Y. Kuniyoshi, M. Inaba, and H. Inoue (1994), "Learning by Watching: Extracting Reusable Task Knowledge from Visual Observation of Human Performance", in *IEEE Transactions on Robotics and Automation,* Vol. 10, No. 6, Dec. 1994

Maja J Matarić and Marc Pomplun (1997), "What do People Look at When Watching Human Movement?", Brandeis University Computer Science Technical Report CS-97-194, July 1997. Submitted to Cognitive Brain Research.

Maja J Matarić, Victor Zordan, and Zach Mason (1998), "Movement Control Methods for Complex, Dynamically Simulated Agents: Adonis Dances the Macarena", in *Proceedings, Autonomous Agents '98,* Minneapolis/St. Paul, May 10-13, 1998.

Maja J Matarić, Matthew Williamson, John Demiris, and Aswath Mohan (1998), "Behavior-Based Primitives for Articulated Control", in *Proceedings, Simulation of Adaptive Behavior '98,* Zurich, Switzerland, August 17-21, 1998.

Andrew Meltzoff, 1996, "The Human Infant as Imitative Generalist: A 20-Year Progress Report on Infant Imitation with Implications for Comparative Psychology", in *Social Learning in Animals: The Roots of Culture*, Heyes C.M. and Galef B.G. (eds), New York: Academic Press.

Ferdinando A. Mussa-Ivaldi, Simon F. Giszter, and Emilio Bizzi (1994), "Linear combinations of primitives in vertebrate motor control", in *Proceedings of the National Academy of Sciences*, vol. 91, pp. 7534–7538, August 1994.

G. Rizzolatti, L. Fadiga, V. Gallese, L. Fogassi, 1996, "Premotor cortex and the recognition of motor actions", in *Cognitive Brain Research,* 3, 131-141.

SD/Fast Users' Manual, 1990, Technical Report, Symbolic Dynamics Inc.

Viviani and Stucchi, 1992, "Motor-perceptual interactions", in *Tutorials in Motor Behaviour II*, G. E. Stelmach and J. Requin (eds), Elsevier publishers, pp.229-248.