Available online at www.sciencedirect.com

SCIENCE DIRECT®

Robotics and Autonomous Systems 54 (2006) 370–384

Robotics and Autonomous Systems

www.elsevier.com/locate/robot

Discriminative and adaptive imitation in uni-manual and bi-manual tasks

Aude G. Billard*, Sylvain Calinon, Florent Guenter

Learning Algorithms and Systems Laboratory - LASA, School of Engineering - STI, Ecole Polytechnique Federale de Lausanne - EPFL, EPFL-STI-I2S-LASA, Station 9, 1015 Lausanne, Switzerland

Available online 29 March 2006

Abstract

This paper addresses the problems of *what to imitate* and *how to imitate* in simple uni and bi-manual manipulatory tasks. To solve the *what to imitate* issue, we use a probabilistic method, based on Hidden Markov Models (HMM), to extract the relative importance of reproducing either the gesture or the specific hand path in a given task. This allows us to determine a metric of imitation performance. To solve the *how to imitate* issue, we compute the trajectory that optimizes the metric given the constraints of the robot's body. We validate the methods using a series of experiments where a human demonstrator uses kinesthetics in order to teach a robot how to manipulate simple objects.

Keywords: Discriminative imitation; Learning; Humanoid robots; Kinesthetic teaching; HMM; Lagrange optimization

1. Introduction

Recent advances in Robot Programming by Demonstration RbD, also referred to as *Learning by Imitation*, have identified a number of key issues that need to be solved in order to ensure a generic approach to the transfer of skills between various agents and situations [17,18]. These have been formulated into a set of generic questions, namely *what to imitate*, *how to imitate*, *when to imitate*, and *who to imitate*. These questions were formulated in response to the large body of work on RbD that emphasized ad hoc solutions to sequencing and decomposing complex tasks into *known* sets of actions, performable by both the demonstrator and the imitator, see e.g. [4,12,15,23,26,27]. In contrast to these other works, the above four questions and their solutions aim at being generic, in that they make no assumptions on the type of skills that can be transmitted. The drawback of such a generic approach is that it has yet to show how the methods will perform when scaled up from acquiring basic skills to complex sequences.

In our previous work we have addressed the *what to imitate* question by developing a general architecture to extract the relevant features of a given task. The method that we used relied on computing the statistical variability of each element of the task; see [9,6]. In this paper, we present an extension of this work which attempts to address the "how to imitate" problem in the control of uni-manual and bi-manual manipulation of objects using a pair of robotic arms, each with four degrees of freedom (DOF). This leads us to tackling more generic issues of motor control, namely that of optimizing the arm controller, given specific constraints. Specifically, we extend the pseudo-inverse optimization method for solving the inverse kinematics, so as to determine the optimal imitation strategy, i.e. the strategy that best satisfies all the constraints of a given task.

The issue of "how to imitate", also referred to as the *correspondence problem* [17], was first addressed very generically in a number of studies with simulated abstract agents acting in a Markov world [1,5]. More recent work by [2, 14] also considered the application of such a system to the control of a robotic arm with two degrees of freedom. The solution to the correspondence problem in the latter works was, however, constrained to a particular arm and did not provide a general solution for robotic arms with an arbitrary number of degrees of freedom. Moreover, in each case, the metric of the task was preset. Here, we present a method that first discovers the metric of the task and then extends the classical inverse kinematics solution to solve generically the correspondence problem for an arbitrary robotic arm.

* Corresponding address: Ecole Polytechnique Federale de Lausanne, School of Engineering, LASA Laboratory, Station 9, 1015 Lausanne, Switzerland. Tel.: +41 21 693 5464; fax: +41 21 693 7807.

E-mail address: aude.billard@epfl.ch (A.G. Billard).

0921-8890/$ - see front matter
doi:10.1016/j.robot.2006.01.007

Next, we will illustrate the key issues at stake and briefly present the approach that we took in solving them.

1.1. What to imitate and how to imitate

When learning a new task by imitation, the robot must first determine the relevant features of the task to be imitated ("what to imitate") and, second, adapt its own motor program to produce an optimal imitation ("how to imitate"). Figs. 1 and 2 illustrate these issues in a simple uni-manual task. Let us first consider the problem of determining the relevant features of a task along with their relative importance.

What to imitate

Consider a planar manipulator with two DOFs, performing two demonstrations of a given task, each time starting with a different joint configuration; see Fig. 1. If one was to record the trajectories of the joints $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$ and of the end-effector $\mathbf{x} = \{x_1, x_2\}$ of the manipulator, one would observe that the first set of variables, i.e. the joint angles, varies significantly, while the second set of variables remains fairly constant from one demonstration to the next. Thus the information conveyed by the hand path would appear to be more reliable than that conveyed by the joints. In other words, the task would appear to put stronger constraints on the hand path than on the joint trajectories. Thus, in order to reproduce the task, one would give more weight to reproducing the hand path than the joint trajectories.

In order to give a measure of the how correctly the action was reproduced, one needs a measure of imitation performance, i.e. a cost function. This cost function must explicitly encapsulate the task constraints, as well as give a measure (metric) of their relative importance.

Let us define $H_1(\boldsymbol{\theta}, \boldsymbol{\theta}')$ and $H_2(\mathbf{x}, \mathbf{x}')$ as measures of the discrepancy between demonstrated and reproduced trajectories of the joints ($\boldsymbol{\theta}$, $\boldsymbol{\theta}'$) and hand path ($\mathbf{x}$, $\mathbf{x}'$), respectively, then, without loss of generality, we can define a global measure of the imitation performance by the weighted sum $H(\boldsymbol{\theta}, \boldsymbol{\theta}', \mathbf{x}, \mathbf{x}') = w_1 H_1(\boldsymbol{\theta}, \boldsymbol{\theta}') + w_2 H_2(\mathbf{x}, \mathbf{x}')$. The weights w_1 and w_2 give a measure of the relative importance of each signal. In the previous example, we would set $w_1 < w_2$ to give more importance to following the path of the end-effector than to reproducing the joint trajectories.

How to imitate

Once we have measured the relative importance of each task feature and have incorporated it into the cost function H, we must determine a trajectory for the joints and end-effector of the imitator's arm that is optimal with respect to the cost function. We note that the problem may not be as straightforward as it seems, since the demonstrator and the imitator may differ significantly in their embodiment (arm length, and number of degrees of freedom for each arm). This is typically the case when transferring information from a human to a robot, even when the robot has a humanoid shape (as is the case in our experiments). This idea is illustrated in Fig. 2, where the two segments of the imitator's manipulator differ in length from those of the demonstrator. If the imitator's arm was allowed to directly replay the joint trajectories of the demonstrator's

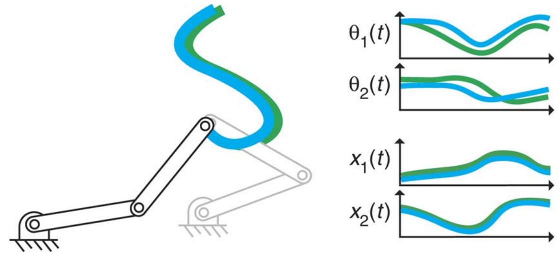

Fig. 1. Illustrative schema of the *what-to-imitate* issue. A two DOF manipulator arm produces two demonstrations of a given task, namely drawing an S figure, and starting with different joint configurations. The path of the end effector given by $\mathbf{x} = \{x_1, x_2\}$ is invariant, while the joint trajectories $\boldsymbol{\theta} = \{\theta_1, \theta_2\}$ vary significantly.

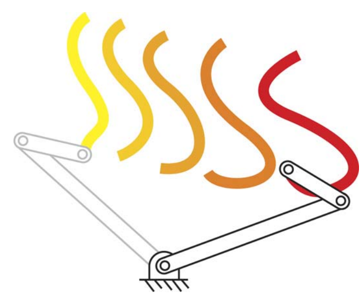

Fig. 2. Illustrative schema of the *how-to-imitate* issue. The manipulator of the imitator generates different alternatives to reproducing the task demonstrated in Fig. 1, from purely satisfying the joint trajectories (left) to satisfying only the hand path (right).

arm, then we would end up with a very different path for the end-effector than the one demonstrated. Conversely, replaying the hand path would result in major differences in the joint trajectories. Fig. 2 illustrates the effect of generating different alternatives, from purely satisfying the joint trajectories (left) to satisfying only the hand path (right).

Minimizing the cost function determines a trade-off between accurately reproducing either the hand path or the joint trajectories. Note that the cost function may have several minima; in other words, there may be several solutions to the problem. we note also that the hand path and the joint trajectories are not independent variables, but are related to each other through a forward and inverse kinematics function, which is what we will discuss next.

1.2. The inverse kinematics problem

In classical control theory, *inverse kinematics* usually refers to the inverse computation required to determine the position of each of the robot's joints for a given location of the robot's end-effector (usually its arm). If we define the coordinates of a manipulator as the n-dimensional vector of joint angles $\vec{\theta}$ and the position of the m-dimensional vector $\vec{x}$, then the forward kinematics are given by: $\vec{x} = f(\vec{\theta})$, while the inverse kinematics are given by

$$\vec{\theta} = f^{-1}(\vec{x}) \tag{1}$$

where f is a continuous function $\in \mathbb{R}$.