Data-based Iterative Human-in-the-loop Robot-Learning for Output Tracking

Rahul B. Warrier * Santosh Devasia **

* Mechanical Engineering Department, University of Washington, Seattle, WA 98195 USA (e-mail: warrierr@uw.edu).
** Mechanical Engineering Department, University of Washington, Seattle, WA 98195 USA (e-mail: devasia@uw.edu)

Abstract: This article develops a data-based approach for improved iterative robot-learning for output-tracking from novice human-in-the-loop demonstrations. While nominal human-response models can be used to improve iterative learning, the convergence can be slow due to variations in each human operator. The major contribution of this article is to use data acquired during iterative learning to learn the unknown human intent as well as the human-response model, and thereby, improve convergence when learning future trajectories. The proposed method is applied to a robot arm, and results indicate both an increase in the range of frequencies where tracking is achieved (from 0.2 Hz to 0.5 Hz) and an increase of 103% in the tracking error reduction for the same number of iterations.

© 2017, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Co-Learning and self-learning; Modeling of human performance; Telerobotics

1. INTRODUCTION

Control systems that involve interactions between human operators and autonomous controllers are referred to broadly as human-in-the-loop control systems, e.g., fly-by-wire aircraft control systems, autonomous driver assistance in cars/trucks, and teleoperated medical devices used in surgeries. Conventionally, autonomous controllers in these systems were developed specific to the task, and required hard-coding the task objectives before deployment. Ideally, more general controllers are desired that can learn how to do specific tasks from the human operator in the loop. In this regard, Learning from Demonstration (LfD), (Argall et al. (2009)), has been identified as a potential solution, wherein the required task is demonstrated by a human operator, which forms the basis for learning by an autonomous controller.

Traditionally, LfD applications assumed that the human demonstrator is an expert model to be emulated (Aghasadeghi et al. (2013)). For example, Dynamic Movement Primitives (DMPs) is a framework that encodes the demonstrated trajectory using linear ordinary differential equations with a learnable nonlinear term, see review by Ijspeert et al. (2013). In this case, there is no direct interaction between the human teacher and the learning robot controller, i.e., human and robot learning do not happen simultaneously. However, recent work has focussed on including the human teacher “in-the-loop” wherein the human teacher is no longer considered as an expert model to be emulated by the robot controller, but rather as an active member in a joint problem-solving team, e.g., see Fong et al. (2003). In this way, even novice human operators can interactively demonstrate the intended task, receiving increasing level of assistance from the autonomous learning controller, leading to faster and more intuitive skill transfer.

With respect to output tracking tasks, it was shown in previous work that the intent of the novice human-in-the-loop demonstrator, i.e., the desired trajectory (when it is unknown) can be inferred by inverting known human feedback response models, e.g., see Warrier and Devasia (2016a,c). The inferred trajectory was used to develop an inversion-based iterative learning control algorithm, where modeling error leads to bounds on the convergence of the iterative scheme. Thus, to ensure convergence, the human models needed to be sufficiently accurate at the required frequencies of interest. Known nominal models of human feedback response, such as the pilot models introduced by McRuer and Jex (1967), are only accurate at low frequencies, restricting their applicability. The major contribution of this article is the development of a model update algorithm that uses data acquired during iterative learning and the learnt desired output trajectory to improve the inference of the unknown human intent, enabling the robot controller to effectively learn to track reference trajectories with higher frequency components than was possible using nominal human feedback models. The proposed method was validated using a human-in-the-loop output tracking task using a robot arm. Experimental results show that the tracking performance at higher target frequencies is improved using the proposed method (about 65% reduction in tracking error compared to manual control) as compared to using nominal human models (about 32% reduction in tracking error compared to manual control), i.e., 103% increase in tracking error reduction.

The rest of the article is organized as follows: Section 2 presents the mathematical formulation of the learning problem, with a brief introduction to the inversion-based learning algorithm developed previously. The proposed data-based model-update algorithm is presented in Section 3. Section 4 provides details on the experimental setup used to validate the proposed algorithm, followed by a
discussion of the results. Conclusions and future work are discussed in Section 5.

2. PROBLEM FORMULATION

2.1 Problem setup

The objective of the task is to achieve the intended goal trajectory, \( y_d \), as the output of the controlled system, \( G \), i.e., \( y = y_d \), as depicted in Fig. 1, where both the human operator (whose response dynamics are denoted by \( G_H \)) and the robot controller provide inputs \( u_h \) and \( u_c \), respectively. In general, the human input, \( u_h \), is a combination of three terms (Wasicko et al. (1966)), described in frequency domain by,

\[
\begin{align*}
\hat{u}_h(\omega) &= G_{H,y_d}(\omega)\tilde{y}_d(\omega) + G_{H,e}(\omega)e(\omega) + G_{H,y}(\omega)y(\omega), \\
&= u_{h,f}(\omega) + u_{h,b}(\omega) + u_{h,y}(\omega)
\end{align*}
\]

(1)

where,

\[
\begin{align*}
G_{H,y_d}(\omega) &= u_{h,f}(\omega)/\tilde{y}_d(\omega) = K_1(\omega)H_{y_d}(\omega), \\
G_{H,e}(\omega) &= u_{h,b}(\omega)/e(\omega) = K_1(\omega)H_e(\omega), \\
G_{H,y}(\omega) &= u_{h,y}(\omega)/y(\omega) = K_1(\omega)H_y(\omega),
\end{align*}
\]

(2)-(4)

where \( K_1 \) represents the human-machine interface, \( H_{y_d}, H_e, H_y \) are the feedforward, feedback and internal response transfer functions of the human operator’s dynamics, respectively. The resulting system output, \( y \), may be described in frequency domain by,

\[
y(\omega) = G(\omega)(u_e(\omega) + u_h(\omega)) = G_{fb}(\omega)\hat{y}_d(\omega) + G_{ff}(\omega)u_e(\omega)
\]

(5)

where the feedback, \( G_{fb} \) and feedforward, \( G_{ff} \) transfer functions are given by,

\[
\begin{align*}
G_{fb}(\omega) &= \frac{(G_{H,y_d}(\omega) + G_{H,e}(\omega)G(\omega))}{1 + (G_{H,x}(\omega) - G_{H,y}(\omega))G(\omega)} = G_{H,1}(\omega)G(\omega), \\
G_{ff}(\omega) &= \frac{G(\omega)}{1 + (G_{H,x}(\omega) - G_{H,y}(\omega))G(\omega)} = \frac{G(\omega)}{1 + G_{H,2}(\omega)G(\omega)},
\end{align*}
\]

(6)-(7)

respectively, all expressed as frequency response functions, i.e.,

\[
G(\omega) = G(s)|_{s=j\omega}, \quad \text{where} \ j = \sqrt{-1}.
\]

(8)

Note that the feedback and feedforward terms in Eqs. (6) and (7) are related as,

\[
G_{fb}(\omega) = G_{H,1}(\omega)G_{ff}(\omega),
\]

(9)

where,

\[
\begin{align*}
G_{H,1}(\omega) &= G_{H,e}(\omega) + G_{H,y}(\omega), \\
G_{H,2}(\omega) &= G_{H,e}(\omega) - G_{H,y}(\omega).
\end{align*}
\]

(10)

Using (10) in (1), the human input, \( u_h \) may be expressed as,

\[
u_h(\omega) = G_{H,1}(\omega)\tilde{y}_d(\omega) - G_{H,2}(\omega)y(\omega).
\]

(11)

Problem statement: Learn the control input, \( u^*_c \) that exactly tracks the desired output trajectory, \( y = y_d \) when the robot controller is subject to the following two constraints:

Constraint 1 (access to desired output): The desired output trajectory, \( y_d \) is not available to the robot controller, but it has direct access to the control input from the human operator, \( u_h \), and the system output, \( y \).

Constraint 2 (modeling uncertainty): The robot controller does not have exact knowledge of the feedback transfer function, \( G_{fb} \), but has access to a model, \( \hat{G}_{fb} \). The control input, \( u = u_h + u_c = u^*_c \) required for exact tracking, \( y = y_d \) for the controlled system \( G \) is given by (5) as,

\[
u^*_c(\omega) = u^*_h(\omega) + u^*_c(\omega) = G^{-1}(\omega)\tilde{y}_d(\omega).
\]

(12)

If the human user is an expert at the task, then the human input \( u_h \) can lead to perfect tracking, i.e., \( y = y_d \). This implies, the robot controller can directly use the human input, \( u_h \approx u^*_c \) (no longer requiring the human input) to achieve perfect tracking. This is the basis of imitation learning approaches, for example, (Schaal (1999)). However, if the human user is not an expert or if the system is complex, then \( u_h \neq u^*_c \) resulting in reduced tracking accuracy. If, instead, the robot controller has access to the desired trajectory, \( y_d \) and the system model \( G \), then it does not require the human input, but can choose the control input, \( u_c = u^*_c = G^{-1}y_d \), to achieve perfect tracking. Moreover, Iterative Learning Control (ILC) may be used in the presence of modeling errors (Arimoto et al. (1984)). But, none of this is possible if the desired trajectory, \( y_d \) is not available to the robot controller, as specified by Constraint 1.

Thus, the human operator’s intended goal, \( y_d \) must be determined using the available information, i.e., the system output, \( y \) and the human input \( u_h \).

3. METHOD

3.1 Model-based human intent estimation

Human intent in the current setup is the desired output trajectory, \( y = y_d \), that the human operator is trying to track with the system, \( G \), as shown in Fig. 1. Then, the intended goal, \( y_d \) may be inferred from (11) as,

\[
y(\omega) = G^{-1}_{H,1}(\omega)[u_h(\omega) + G_{H,2}(\omega)y(\omega)].
\]

(13)

forming the intent estimate, \( \hat{y}_d \) given by,

\[
\hat{y}_d(\omega) = \hat{G}_{H,1}(\omega)[u_h(\omega) + \hat{G}_{H,2}(\omega)y(\omega)],
\]

(14)

where,

\[
\begin{align*}
\hat{G}_{H,1}(\omega) &= \hat{G}_{H,e}(\omega) + \hat{G}_{H,y}(\omega), \\
\hat{G}_{H,2}(\omega) &= \hat{G}_{H,e}(\omega) - \hat{G}_{H,y}(\omega),
\end{align*}
\]

(15)

where \( \hat{G}_{H,y}, \hat{G}_{H,e}, \text{ and } \hat{G}_{H,y} \) are known models of the human response transfer functions defined in Eqs. (2)-(4).

3.2 Learning the controller input using estimated intent

Consider the system output at the \( k \)th iteration, \( y_k \), which is obtained from (5) as,

\[
y_k(\omega) = G_{fb}(\omega)\hat{y}_d(\omega) + G_{ff}(\omega)u_{c,k}(\omega),
\]

(16)
دریافت فوری متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات