



HAL
open science

Optimizing Human-Robot Interactions through Differential Game Control

Abdelwaheb Hafs, Dorian Verdel, Waldez Gomes, Olivier Bruneau, Bastien
Berret

► **To cite this version:**

Abdelwaheb Hafs, Dorian Verdel, Waldez Gomes, Olivier Bruneau, Bastien Berret. Optimizing Human-Robot Interactions through Differential Game Control. Workshop on Multilimb Coordination in Human Neuroscience and Robotics: Classical and Learning Perspectives, IROS, Oct 2023, Detroit (MI), United States. hal-04258166

HAL Id: hal-04258166

<https://hal.science/hal-04258166>

Submitted on 25 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimizing Human-Robot Interactions through Differential Game Control

Abdelwaheb Hafis , Dorian Verdel , Waldez Gomes , Olivier Bruneau , Bastien Berret 

I. INTRODUCTION

Contact robots have been increasingly used to physically assist humans in industrial or medical contexts. To fully exploit the potential of these devices, it is essential to improve the intuitiveness of interaction control [1]. One way of doing it is estimating the human motion intent and sharing the task effort [2]. Here, we present a method for co-adaptive shared control wherein a human and a robot adapt their respective control towards each other, particularly when the intended motion is known over a finite time horizon.

An appealing framework to formalize interaction control is the theory of differential games (DG) [2]. By estimating the human cost function, we could theoretically predict the upcoming human control input and adapt the robot's cost following DG optimality principles [3]. Recent studies have illustrated the potential of such game-theoretic approaches for sharing effort during physical human-robot interaction. For example, Li et al. successfully used non-cooperative DG to optimally assist humans during reaching movements [4]. Musić et al. generalized this approach to the tracking of arbitrary trajectories [1]. However, these works used an infinite time horizon in practical implementations, which is unlikely in real scenarios. We propose a solution to this limitation to enable effort sharing in a finite-horizon context.

II. METHODS

A. Dyadic system model

We consider the physical interaction between a robot and a human during a trajectory-tracking task, the reference trajectory being known *a priori*. We focus on the problem of effort sharing. To this aim, we assume that the two agents continuously play a non-cooperative DG over a finite time horizon. We assume that the human cost function summarizes all human motor behavior, so it must be inferred during task execution. We present below a method to estimate its parameters, which is an instance of inverse DG problem [5].

Let us assume that the interaction dynamics can be described by the following control system:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t), t) \quad (1)$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ and $\mathbf{u}(t) \in \mathbb{R}^m$ are the state and control of the system at time t , respectively. The control \mathbf{u} includes the human and robot inputs.

AH, WG, DV and BB are with the CIAMS, Université Paris-Saclay, 91405 Orsay, France. Corresponding e-mail: abdelwaheb-hafs@universite-paris-saclay.fr.

OB is with the LURPA, ENS Paris-Saclay, 91190 Gif-sur-Yvette, France.

This work is supported by the French National Agency for Research (grant ANR-19-CE33-0009).

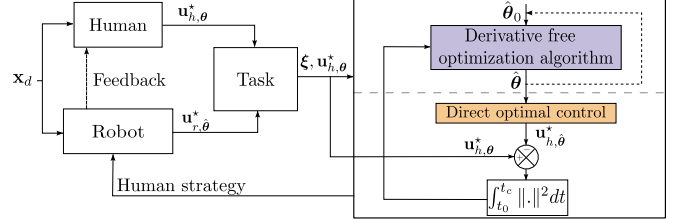


Fig. 1: Control and cost estimation workflow. \mathbf{x}_d and $\boldsymbol{\xi}$ are the desired state and the state deviation respectively, $\mathbf{u}_{h,\theta}^*$ is the (optimal) human control, $\hat{\boldsymbol{\theta}}$ and $\mathbf{u}_{h,\hat{\theta}}^*$ are the estimated human cost parameters and corresponding (optimal) human control, $\boldsymbol{\theta}_0$ is the initial guess of the parameters of the cost function.

We also assume that there exists a desired trajectory $\mathbf{x}_d(t)$ solution of (1) for $t \in [0, T]$ with an associated control $\mathbf{u}_d(t)$, where T is the total duration of the tracking task. By linearizing the dynamics (1) around this reference trajectory-control pair, the system can be described by the following affine dynamics:

$$\dot{\boldsymbol{\xi}}(t) = \mathbf{A}(t)\boldsymbol{\xi}(t) + \mathbf{B}(t)\mathbf{u}(t) - \mathbf{B}(t)\mathbf{u}_d(t) \quad (2)$$

where $\mathbf{A}(t)$ and $\mathbf{B}(t)$ are the state and input Jacobian matrices, $\boldsymbol{\xi}(t) = \mathbf{x}(t) - \mathbf{x}_d(t)$ is the state deviation, and $\mathbf{u}(t) = \mathbf{u}_r(t) + \mathbf{u}_h(t)$ is the input. The subscript r refers to the robot and h refers to the human throughout the paper.

Assuming quadratic objectives for each agent, this interaction dynamics fits the class of tractable DG problems (i.e., an affine-quadratic DG for which necessary and sufficient optimality conditions of a feedback Nash equilibrium solution are available [3]).

B. Human-robot differential game

Therefore, we assume that both agents minimize a quadratic cost as defined below:

$$J_r = \int_{t_c}^{t_c+\Delta} \boldsymbol{\xi}(t)^\top \mathbf{Q}_r \boldsymbol{\xi}(t) + \mathbf{u}_r(t)^\top \mathbf{R}_r \mathbf{u}_r(t) + \mathbf{u}_h(t)^\top \mathbf{R}_{rh} \mathbf{u}_h(t) dt \quad (3)$$

$$J_h = \int_{t_c}^{t_c+\Delta} \boldsymbol{\xi}(t)^\top \mathbf{Q}_h \boldsymbol{\xi}(t) + \mathbf{u}_h(t)^\top \mathbf{R}_h \mathbf{u}_h(t) dt \quad (4)$$

where t_c is the current (initial) time and Δ is the finite time horizon used for the DG. The matrices \mathbf{Q}_r and $\mathbf{Q}_h \in \mathbb{R}^{n \times n}$ are positive semi-definite (assumed to be diagonal here) that define the weights associated with tracking error. \mathbf{R}_r and $\mathbf{R}_h \in \mathbb{R}^{m \times m}$ are positive definite matrices that define the weights associated with the effort of the robot and human

respectively. Finally, $\mathbf{R}_{rh} \in \mathbb{R}^{m \times m}$ is a positive semi-definite matrix that specifies that the robot is designed to assist the human, and thus aims at reducing the human effort. In this situation, a Nash equilibrium exists, and the optimal control of each agent can be obtained by solving a set of ordinary differential equations backwards in time [3]. The actual optimal controls of the robot and human will be denoted by \mathbf{u}_r^* and \mathbf{u}_h^* respectively. The latter is unknown but could be estimated by identifying the parameters of the human cost. These parameters are the weights defined in \mathbf{Q}_h and \mathbf{R}_h and then collected in a vector θ . The goal is to find an estimate $\hat{\theta}$ of θ .

C. Bi-level estimation method

We used a bi-level approach to estimate θ by solving the above lower-level DG problem (Eq. 3 and Eq. 4) together with the following upper-level problem:

$$\min_{\hat{\theta}} \int_{t_c - \Delta}^{t_c} \|\mathbf{u}_{h,\theta}^*(t) - \mathbf{u}_{h,\hat{\theta}}^*(t)\|^2 dt \quad (5)$$

where the control $\mathbf{u}_{h,\hat{\theta}}^*$ is the optimal control of the non-cooperative DG with the estimated cost parameters $\hat{\theta}$ and $\mathbf{u}_{h,\theta}^*$ is the optimal human control with the true cost parameters θ , over the time horizon Δ (called estimation horizon). The main difficulty with this bi-level approach is that it may be computationally intensive, which can limit its application to real-time control. However, it is possible to use only a few iterations of the optimization to improve $\hat{\theta}$, which can be used as the initial guess for the next time step. Note that the method is applied in a receding horizon fashion, so the problem at each time step of the robot control loop will be different as new observations come in.

D. Simulation

To test the control and estimation method presented above, we simulated a trajectory tracking task involving the human forearm connected to a 1-dof robot. The task was to track a trajectory in the vertical plane (with simulated gravitational forces). We simulated a human with a height of 1.70 m and weight of 70 kg using anthropometric values in [6]. The human was assumed to know the robot control strategy, while the robot estimated the human control. The effort cost parameters were set to 1 because inputs are scalars and of the same order of magnitude. We assumed that the error cost matrices were $\mathbf{Q}_r = \mathbf{Q}_h = \text{diag}([20, 5])$, so that the human cost parameters are represented in one vector of 2 parameters $\theta = [\theta_1, \theta_2]^T = [20, 5]^T$. The task duration was set to $T = 30$ s. We included white noise at an intensity of 0.1 Nm to simulate variability in human input \mathbf{u}_h^* . The finite control horizon and the estimation horizon were set to $\Delta = 1$ s. Finally the human cost matrix switches to $\mathbf{Q}_h = \text{diag}([30, 10])$ in the middle of the task to simulate a change in the human contribution to the task.

III. RESULTS AND DISCUSSION

Fig. 2 illustrates the simulated and estimated human cost parameters with their corresponding control inputs. The

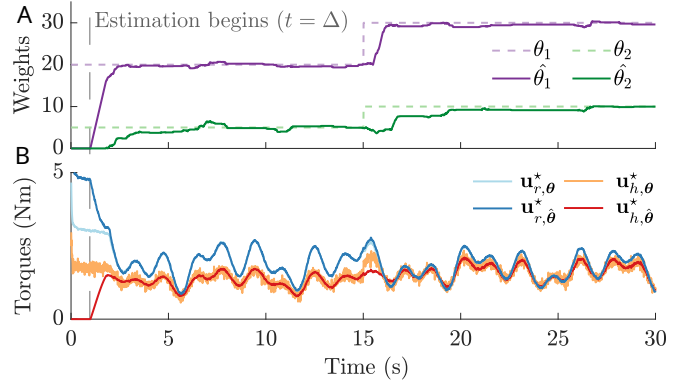


Fig. 2: Estimation of the human control strategy. The simulated robot mass, inertia and damping were $M_r = 1$ kg, $I_r = 0.0625$ kg.m² and $D_r = 0.5$ kg.m².s⁻¹, and the human’s forearm mass, inertia, damping and stiffness were $M_h = 1.561$ kg, $I_h = 0.0456$ kg.m², $D_h = 0.5$ kg.m².s⁻¹ and $K_h = 1$ kg.m².s⁻². Distance to the centers of gravity was set to $L = 0.202$ m. A: The true simulated human cost parameters (position weight θ_1 and velocity weight θ_2) are in faded-dashed lines, and the estimated parameters are in plain lines. B: The true Nash equilibrium torques are in faded lines, and the torques using the estimation method are in bold lines.

initial guess $\hat{\theta}_0$ was set to $\mathbf{0}$, the results showed that our approach estimates the parameters θ quite rapidly. The root mean square (RMS) errors between the recovered and actual parameters were $[2.522 \quad 1.594]^T$. We observed that the estimated human cost parameters generated well the theoretical Nash equilibrium for optimal shared control of the task with corresponding RMS errors of 0.186 Nm and 0.145 Nm for the human and robot respectively.

IV. CONCLUSIONS

This paper introduced an inverse DG approach for co-adaptive shared control applicable to non-linear dynamical systems when a desired state/control trajectory is available over a relatively short time horizon. A bi-level optimization approach for recovering the human cost parameters was sketched and tested in simulation. Future work will consider this approach in a real human-robot experiment and analyze the resulting co-adaptation.

REFERENCES

- [1] S. Musić and S. Hirche, “Haptic shared control for human-robot collaboration: A game-theoretical approach,” *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 10216–10222.
- [2] Y. Li, A. Sena, Z. Wang, X. Xing, J. Babič, E. v. Asseldonk, and E. Burdet, “A review on interaction control for contact robots through intent detection,” *Progress in Biomedical Engineering*, vol. 4, no. 3, p. 032004, Aug 2022.
- [3] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory, 2nd Edition*, ser. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.
- [4] Y. Li, G. Carboni, F. Gonzalez, D. Campolo, and E. Burdet, “Differential game theory for versatile physical human–robot interaction,” *Nature Machine Intelligence*, vol. 1, no. 1, pp. 36–43.
- [5] T. L. Molloy, J. Inga Charaja, S. Hohmann, and T. Perez, *Inverse Optimal Control and Inverse Noncooperative Dynamic Game Theory: A Minimum-Principle Approach*, ser. Communications and Control Engineering. Springer International Publishing.
- [6] D. A. Winter, *Biomechanics and motor control of human gait: normal, elderly and pathological - 2nd edition*, 1991, vol. Ed2.