ScienceDirect
IFAC PapersOnLine 53-2 (2020) 6670–6677

Optimal Control and Inverse Optimal Control with Continuous Updating for Human Behavior Modeling

Ovanes Petrosian * **Jairo Inga** ** **Ildus Kuchkarov** *
Michael Flad ** **Sören Hohmann** **

* *Faculty of Applied Mathematics and Control Processes, Saint Petersburg State University, Saint Petersburg, Russia, (e-mail: petrosian.ovanes@yandex.ru).*
** *Institute of Control Systems, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany*

Abstract: The theory of optimal control has received considerable attention to model motion behavior or decision making of humans. Most approaches are based on a fixed (or infinite) time horizon which implies that all information is available at the beginning of the time interval. Nevertheless, it is reasonable to believe that the human uses information defined by a continuously moving information horizon at each time instant and adapts accordingly. Therefore, in this paper, we propose an optimal feedback control approach based on the paradigm of continuous updating. The model parameters which define individual human behavior consist of the cost function parameters and the length of the information horizon, which can be identified via a corresponding inverse optimal control approach. We show the applicability of the approach with simulations of a potential application example of human behavior identification from the point of view of a driving assistance system.

1. INTRODUCTION

Understanding how humans move or control a system is an important issue in human-machine interaction, e.g. in shared control where a human and an automatic controller influence a dynamic system simultaneously. The knowledge of human behavior can be used to adapt the controller according to a particular human partner. The theory of optimal feedback control arised in the neuroscience community as a promising approach to model human motor behavior (Todorov (2004)). The last two decades have therefore seen a growing interest in optimal control theory to model all kinds of human movement including reach-to-grasp (El-Hussieny et al. (2016)), saccadic eye movements (El-Hussieny and Ryu (2018)) and locomotion (Mombaur et al. (2010)). This has also motivated the use of optimal control to model how a human controls or manipulates a dynamic system alone (Priess et al. (2015); Inga et al. (2017)) or in cooperation with an assistance system (Inga et al. (2018)). For this purpose, the model parameters are determined out of measured data by solving inverse optimal control problems. This problem has also attracted much attention from the control engineering community (see e.g. Johnson et al. (2013); Pauwels et al. (2014); Jean and Maslovskaya (2018)).

In all the above modeling approaches for human behavior, it is assumed that the human has all information about the motion equations and objective function at the beginning of the process and for the complete time interval which may be finite (e.g. Aghasadeghi and Bretl (2014)) or infinite (e.g. Priess et al. (2015); El-Hussieny et al. (2016)). Nevertheless, most of real-life control processes evolve continuously in time, and the human may not have all information about the process at the initial time instant. For example, in an automotive application, the driver only has local information about the road curvature or any obstacles which may force a lane change. Hence, it is questionable whether approaches based on classical optimal control theory reflect human decision making adequately. Therefore, we conjecture that a human continuously receives updated information and adapts his behavior continuously to the new situation and hence, it is important to include this characteristic in an identification procedure of a human behavior model.

Some literature exists where it is assumed that humans have a preview time up to which they have knowledge of the systems states or plan their actions, especially in the field of human driver modeling for automotive applications, see e.g. the work in Gray et al. (2013); Flad et al. (2014); Inga et al. (2015); Zhang et al. (2019). Nevertheless, this preview time is usually set manually and not identified out of measured data. Understanding the length of the information horizon with an identification approach can be beneficial to better understand human decision making.

The idea of an information horizon has also been explored in the papers of Gromova and O.L. (2016), Petrosian (2016), Petrosian and Barabanov (2017), Petrosian et al. (2018), Yeung and Petrosian (2017), Petrosian and Gromova (2018), Petrosian and Kuchkarov (2019), Petrosian et al. (2019). In these papers it is assumed that the information about motion equations and payoff functions is updated in discrete time instants. In the papers Petrosian and Tur (2019), Kuchkarov and Petrosian (2019)

10.1016/j.ifacol.2020.12.089

information about the process evolves continuously in time. In the paper Petrosian and Tur (2019), a system of Hamilton-Jacobi-Bellman equations is derived for solving a Nash dynamic game with continuous updating. Similarly, Kuchkarov and Petrosian (2019) explore the class of linear-quadratic differential games with continuous updating and derive the explicit form of the Nash equilibrium.

The class of control problems with dynamic updating has some similarities with Model Predictive Control (MPC) theory which is worked out within the framework of numerical optimal control (Goodwin et al. (2005), Kwon and Han (2005), Rawlings and Mayne (2009), Wang (2005)) and which has also been used as a human behavior model in Ramadan et al. (2016). In the MPC approach, the current control action is achieved by solving a finite-horizon open-loop optimal control problem at each sampling instant.For linear systems there exists a solution in explicit form (Hempel et al. (2015), Bemporad et al. (2002)). However, in general, the MPC approach demands the solution of several optimization problems. Another related series of papers corresponds to the class of stabilizing control (Mayne and Michalska (1990), Kwon et al. (1982), Kwon and Pearson (1977), Shaw (1979)). While these techniques to determine optimal control are close to the ones presented in this paper, their considered problem is usually assumed to have a strict terminal contraint for the states in each horizon.

In this paper, we present a human behavior modeling approach based on the paradigm of continuous updating. This solves the problem of modeling human behavior when information about the process updates continuously in time. This means that the human

- has information about motion equations and objective function only on $[t, t + \overline{T}]$, where $\overline{T}$ is the information horizon and t is the current time instant.
- receives updated information as time $t \in [t_0, +\infty)$ evolves.

Furthermore, we leverage explicit solutions of a linear-quadratic optimal control problem with continuous updating to propose a corresponding inverse optimal control approach, where not only the human cost functions parameters are identified, but also the length of the information horizon $\overline{T}$.

The paper is structured as follows. In Section 2, a description of the initial optimal control problem and corresponding optimal control problem with continuous updating is presented. In Section 3, the explicit form of the solution of the optimal control problem with continuous updating is given for the class of linear-quadratic control problems. Then, we present in Section 4 the inverse optimal control approach with continuous updating, where both cost function parameters and the length of the information horizon are identified. Afterwards, we show in Section 5 simulation results of the proposed modeling approach based on continuous updating with parameter identification. Finally, we draw conclusions in Section 6.

2. OPTIMAL CONTROL WITH CONTINUOUS UPDATING

In this section, we present our results concerning optimal control with continuous updating. We first present the classical optimal control problem before showing our continuous updating approach.

2.1 Initial Optimal Control Problem

Consider the optimal control problem defined on the interval $[t_0, T]$:

$$J(x_0, t_0; u) = \int_{t_0}^{T} g[t, x(t), u(t, x)]dt \ \rightarrow \min_{u} \tag{1}$$

subject to

$$\begin{aligned} &\dot{x}(t) = f(t, x, u), \\ &x(t_0) = x_0, \\ &x \in \text{comp}\mathbb{R}^n, u = u(t, x) \in U \subset \text{comp}\mathbb{R}^m, t \in [t_0, T], \end{aligned} \tag{2}$$

where $\text{comp}\mathbb{R}^m$ is a compact set in an m-dimensional space of real numbers, $g[t, x(t), u(t, x(t))]$, $f(t, x, u)$ are integrable functions, $x(t) \in \mathbb{R}^n$ is the solution of the Cauchy problem (2) with fixed $u(t, x) \in \mathbb{R}^m$. The control $u(t, x) \in \mathbb{R}^m$ is called admissible if the problem (2) has a unique and continuous solution.

2.2 Problem Formulation for Optimal Control with Continuous Updating

Using the initial optimal control problem defined on the closed time interval $[t_0, T]$, we construct the corresponding optimal control problem with continuous updating.

Consider the following optimal control problem defined on the interval $[t, t + \overline{T}]$, where $0 < \overline{T} < +\infty$:

$$J(x, t; u^t) = \int_{t}^{t+\overline{T}} g[s, x_t(s), u^t(s, x_t)]ds \ \rightarrow \min_{u^t} \tag{3}$$

subject to

$$\begin{aligned} &\dot{x}_t(s) = f(s, x_t, u^t), \\ &x_t(t) = x, \\ &x_t \in \text{comp}\mathbb{R}^n, u^t = u^t(s, x_t) \in U \subset \text{comp}\mathbb{R}^m, s \in [t, t + \overline{T}], \end{aligned} \tag{4}$$

where $u^t(s, x_t) \in \mathbb{R}^m$ and $x_t(s) \in \mathbb{R}^n$ are the optimal control and the optimal state trajectories on the interval $[t, t + \overline{T}]$, respectively.

The main characteristic of the optimal control problem with continuous updating is the following:

The current time $t \in [t_0, +\infty)$ evolves continuously and as a result the human continuously obtains new information about motion equation and objective function on the interval $[t, t + \overline{T}]$.

The control $u(t, x)$ in the optimal control problem with continuous updating has the form:

$$u(t, x) = u^t(s, x)|_{s=t}, \ t \in [t_0, +\infty), \tag{5}$$

where $u^t(s, x)$, $s \in [t, t + \overline{T}]$ is the control in the problem defined on the interval $[t, t+\overline{T}]$ and $u^t(s, x)|_{s=t}$ is the part of that control in the first instant $s = t$. The main idea of (5) is that as the current time t evolves information updates, therefore in order to model the behavior of the human it is necessary to consider the control $u^t(s,x)$ only in the points where $s = t$.

The trajectory $x(t)$ in the optimal control problem with continuous updating is determined in accordance with the system dynamics in (2) where $u = u(t,x)$ is the control in (5). We assume that the control with continuous updating obtained using (5) is admissible.

The essential difference between the control problem with continuous updating and classic optimal control problem defined on the closed interval is that the decision maker in the initial problem is guided by the objective that it will eventually receive on the interval $[t_0,T]$, but in the case of a control problem with continuous updating, at the time instant t the system is oriented on the expected objective (3), which is calculated using the information on the interval $[t,t+\overline{T}]$ or the information that the system has at the instant t.

2.3 Optimal Control with Continuous Updating

For the framework of continuously updated information, we use the concept of optimal control in feedback form $u^*(t,x)$. Furthermore, we require that for any fixed current time $t \in [t_0,+\infty)$, $u^*(t,x)$ coincides with the optimal control in the problem specified by (3) and (4), defined on the interval $[t,t+\overline{T}]$ in the instant t.

However, direct application of classical approaches for optimal control in feedback form is not possible. Consider another current time instant $t+\epsilon$, $\epsilon << \overline{T}$, then according to the aforementioned requirement, $u^*(t+\epsilon,x)$ in the instant $t+\epsilon$ must coincide with the optimal control in the problem defined on the interval $[t+\epsilon, t+\epsilon+\overline{T}]$.

Therefore $u^*(t,x)$, $t \in [t_0,+\infty)$ should be defined as an infinite combination of optimal controls on intervals $[t,t+\overline{T}]$ for every current time instant $t \in [t_0,+\infty)$.

In order to construct such controls, we consider a concept of generalized optimal control in feedback form

$$\widetilde{u}^*(t,s,x),\ t \in [t_0,+\infty),\ s \in [t,t+\overline{T}], \tag{6}$$

which we are going to use further for construction of the control $u^*(t,x)$.

Definition 2.1. The control $\widetilde{u}^*(t,s,x)$ is a generalized optimal control in the problem with continuous updating if it is optimal, for any fixed $t \in [t_0,+\infty)$, with respect to the problem given by the cost function (3), and the system dynamics (4).

Using generalized optimal control it is possible to define solution concept for an optimal control problem with continuous updating.

Definition 2.2. The control $u^*(t,x)$ is called optimal control with continuous updating if

$$u^*(t,x) = \widetilde{u}^*(t,s,x)|_{s=t},\ t \in [t_0,+\infty), \tag{7}$$

where $\widetilde{u}^*(t,s,x)$ is the generalized optimal control defined in Definition 2.1.

3. LINEAR QUADRATIC OPTIMAL CONTROL WITH CONTINUOUS UPDATING

In this section, we give solutions for the optimal control problem with continuous updating according to Definition 2.2 for the class of linear-quadratic optimal control.

3.1 Problem Formulation for the Linear-Quadratic Case with Continuous Updating

According to the general problem statement of the optimal control problem with continuous updating the linear quadratic case will have the following form:

Consider the optimal control problem defined on the interval $[t,t+\overline{T}]$, where $0 < \overline{T} < +\infty$:

$$J(x,t;u^t) = \int_t^{t+\overline{T}} {x_t}^T(s) Q x_t(s) + u^t(s,x_t)^T R u^t(s,x_t) ds \ \to \min_{u^t} \tag{8}$$

subject to

$$\begin{aligned} &\dot{x}_t(s) = A x_t(s) + B u^t(s,x_t), \\ &x_t(t) = x, \\ &x_t \in \mathbb{R}^n, u^t = u^t(s,x_t) \in U \subset \text{comp}\mathbb{R}^m, s \in [t,t+\overline{T}]. \end{aligned} \tag{9}$$

The definitions of the generalized optimal control and the optimal control with continuous updating are analogous to Definitions 2.1 and 2.2.

3.2 Optimal Control with Continuous Updating for LQP

Here we present the explicit form of the optimal control with continuous updating using the system of Riccati differential equations.

Theorem 3.1. The linear quadratic control problem with continuous updating defined by (8) and (9) has, for every initial state x_0, a solution if and only if the Riccati differential equation (10) has a symmetric solution $K(\cdot)$ on the interval $[0,1]$:

$$\begin{aligned} &\dot{K}(\tau) = -\overline{T}A^T K(\tau) - \overline{T}K(\tau)A + \overline{T}K(\tau)SK(\tau) - \overline{T}Q, \\ &K(1) = 0, \end{aligned} \tag{10}$$

where $S = BR^{-1}B^T$, Q, R are assumed to be symmetric and R is positive definite.

If the linear quadratic control problem with continuous updating has a solution, then it is unique and the optimal control in feedback form is

$$u^*(t,x) = -R^{-1}B^T K(0)x. \tag{11}$$

Proof. In order to prove the Theorem we introduce the following change of variables

$$\begin{aligned} s &= t + \overline{T}\tau, \\ y_t(\tau) &= x(t+\overline{T}\tau), \\ v_t(\tau,y_t) &= u(t+\overline{T}\tau,x). \end{aligned} \tag{12}$$

By substituting (12) in the motion equations (9) and objective function (8) we obtain

$$\dot{y}_t(\tau) = \overline{T}Ay_t(\tau) + \overline{T}Bv_t(\tau,y_t) \tag{13}$$

and

$$J(y_t, t; v_t) = \int_0^1 \overline{T} y_t^T(\tau) Q_i y_t(\tau) + \overline{T}(v_t(\tau, y_t))^T R v_t(\tau, y_t) d\tau. \quad (14)$$

It is known (see e.g. (Engwerda, 2005, Theorem 5.1)) that the criterion for existence of optimal control in feedback form is the existence of symmetric solution for the system of differential equations (10). The optimal control has the form

$$v_t^*(\tau, y_t) = -R^{-1} B^T K(\tau) y_t. \quad (15)$$

From (12) we have

$$\tau = \frac{s-t}{\overline{T}}.$$

Returning to original variables we obtain the following control

$$u^t(s, x) = -R^{-1} B^T K \left(\frac{s-t}{\overline{T}} \right) x.$$

This control is optimal control in the problem defined on the interval $[t, t+\overline{T}]$ by construction. The equations (13), (14) and solution (15) have the same form for all values t. Then a generalized optimal control in the control problem with continuous updating has the form

$$\widetilde{u}^*(t, s, x) = -R^{-1} B^T K \left(\frac{s-t}{\overline{T}} \right) x. \quad (16)$$

By using (7), we set the optimal control with continuous updating equal to the generalized optimal control (16), for $s = t$:

$$u^*(t, x) = -R^{-1} B^T K(0) x, \ t \in [t_0, +\infty). \quad (17)$$

Thus, control (15) in the subtask exists for any initial value y_t $(t \geq t_0)$, then control (17) in the optimal control problem with continuous updating of information exists for any x_0. ■

Using the form of the optimal control with continuous updating $u^*(t, x)$ given in (11), it is possible to propose a corresponding inverse optimal control problem with continuous updating, which we present in the following.

4. INVERSE OPTIMAL CONTROL PROBLEM WITH CONTINUOUS UPDATING

Suppose the function $g[t, x(t), u(t, x(t))]$ in (1) is parameterized and therefore, in an inverse optimal control problem, this parametrization is unknown and has to be estimated from observed control and state trajectories. Furthermore, the value of the information horizon $\overline{T}$ is also unknown. We denote by θ the set of unknown parameters including the cost function parametrization and the information horizon. The inverse control problem with continuous updating involves the estimation of the unknown parameter set θ based on the observed control $\hat{u}(t)$ and corresponding trajectory $\hat{x}(t)$ on the interval $[t_0, T]$. The objective is to obtain estimated optimal control $u_\theta^*(t, x_\theta^*(t))$ and corresponding trajectory $x_\theta^*(t)$ such that the difference between these model trajectories and the observed ones is minimal.

Therefore, in this paper, the inverse optimal control problem with continuous updating is solved with an approach analogous to the bi-level approach for standard optimal control introduced in Mombaur et al. (2010), where it has successfully been applied with experimental data. The model parameters θ are determined such that the squared error between observed and model trajectories is minimized. This objective is represented by the optimization problem

$$Div = \int_{t_0}^{T} ||x_\theta(t) - \hat{x}(t)||^2 + ||u_\theta(t, x_\theta^*(t)) - \hat{u}(t)||^2 \, \mathrm{d}t \ \rightarrow \min_\theta, \quad (18)$$

where x_θ and u_θ are the optimal control trajectories arising from the minimization of the cost function (1) with parameters θ. Note that in the continuous updating inverse optimal control approach, $u_\theta(t, x_\theta(t))$ is defined as in (7) and $x_\theta(t)$ arises from the corresponding solution of the system dynamics. Therefore, in the following, we consider the linear-quadratic optimal control problem with continuous updating such that we are able to exploit the explicit form given in (11) for the solution of (18).

5. SIMULATION RESULTS

In this section, we illustrate our proposed optimal control approach with continuous updating and how the corresponding solution of the inverse problem can be applied to identify suitable parameters to describe behavior of the driver controlling the lateral dynamics of a vehicle.

5.1 Single-track and Steering Model

The simulated system is a single-track and steering model with six states from Flad et al. (2014). The model is a validated linear approximation of steering behavior in a midsize passenger car.

The system state vector is given by

$$x(t) = \left[\beta(t) \ \psi(t) \ \dot{\psi}(t) \ y(t) \ \delta(t) \ \dot{\delta}(t) \right]^T,$$

where $\beta(t)$ is the side-slip angle, $\psi(t)$ is the yaw angle, $\dot{\psi}(t)$ is the yaw angle velocity , $y(t)$ the lateral distance from middle lane, $\delta(t)$ is steering wheel angle and $\dot{\delta}(t)$ is the steering wheel angle velocity. All states are in SI units. The control $u(t, x) \in \mathbb{R}$ in the system is the applied torque by the human.

The system dynamics have the form:

$$\dot{x}(t) = A\, x(t) + B\, u(t) \quad (19)$$

with $x(t_0) = x_0$ given and

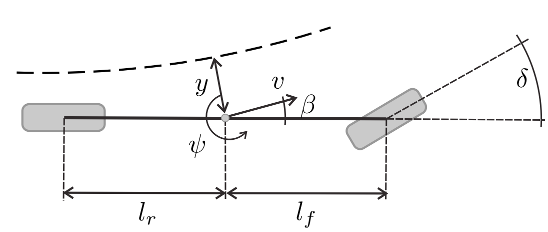

Fig. 1. Geometric relations of the single-track vehicle model.

$$A = \begin{bmatrix} \frac{-C_f - C_r}{Mv} & \frac{C_r l_r - Mv^2 - C_f l_f}{Mv^2} & 0 & 0 & \frac{C_f}{Mvi_s} & 0 \\ \frac{C_r l_r - C_f l_f}{J_z} & \frac{-C_f l_f^2 - C_r l_r^2}{J_z v} & 0 & 0 & \frac{C_f l_f}{J_z i_s} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ v & 0 & v & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{-C_s}{J_s} & \frac{-D_s}{J_s} \end{bmatrix}, \quad (20)$$

$$B = \left[0\ 0\ 0\ 0\ 0\ \frac{1}{J_s} \right]^T.$$

The model geometric relations are depicted schematically in Fig. 1. The system's parameters can be found in Table 1.

Table 1. System constants of the linear single-track steering model of Flad et al. (2014)

Name	Value	Unit	Description
v	20	$\frac{\text{m}}{\text{s}}$	vehicle velocity
l_f	1.5	m	length of center of mass to front wheelbase
l_r	1	m	length of center of mass to rear wheelbase
C_f	137.5	$\frac{\text{kN}}{\text{rad}}$	front tire slip
C_r	137.5	$\frac{\text{kN}}{\text{rad}}$	rear tire slip
M	1500	kg	vehicle mass
J_z	1800	$\text{kg}\,\text{m}^2$	vehicle yaw inertia
D_s	1	$\frac{\text{Nm}\,\text{s}}{\text{rad}}$	damping constant of steering train
C_s	1	$\frac{\text{Nm}}{\text{rad}}$	retraction constant of steering wheel
J_s	0.1	$\frac{\text{Nm}\,\text{s}^2}{\text{rad}}$	inertia of the steering train
i_s	16	–	steering transfer constant

.

Similar to e.g. Cole et al. (2006), we suppose that the objective function of the driver has a quadratic form, therefore the integrand in (1) and (8) has the form:

$$g[t, x(t), u(t,x), \theta] = x^T(t)Qx(t) + u(t,x)^T Ru(t,x). \quad (21)$$

The parameter θ has to be identified out of the generated data by means of the approach in (18). The parameter θ is composed of the elements of the matrices Q and R and the time horizon $\overline{T}$:

$$\theta = \{Q, R, \overline{T}\}. \quad (22)$$

5.2 Optimal Control with Continuous Updating

In order to illustrate the different solutions arising from optimal control with continuous updating, we first define the ground truth parameters of the objective function of the human driver to have the form:

$$Q = diag(1, 10, 1, 40, 10, 1), \quad R = [1]. \quad (23)$$

Furthermore, suppose the initial state at $t_0 = 0$ is

$$x_0 = [0\ 0\ 0\ 1\ 1\ 0.1].$$

We use optimal control with continuous updating, i.e. the results of Theorem 3.1 to generate data sets which simulate measurements of human steering behavior. This was done by solving (10) numerically using the `ode45` solver of MATLAB. The data sets, denoted by D1 to D4, consist of optimal state and control trajectories with $t \in [0, 10]$ and with the parameters for the information horizon

$$\overline{T} = 1,\ \overline{T} = 1.5,\ \overline{T} = 2,\ \overline{T} = 2.5, \quad (24)$$

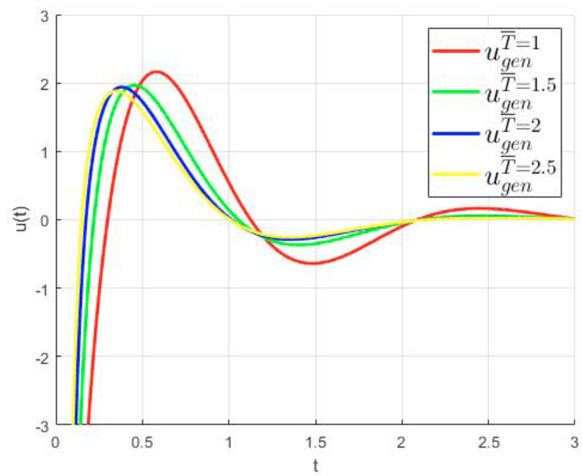

Fig. 2. Optimal controls with continuous updating with different lengths of the information horizon.

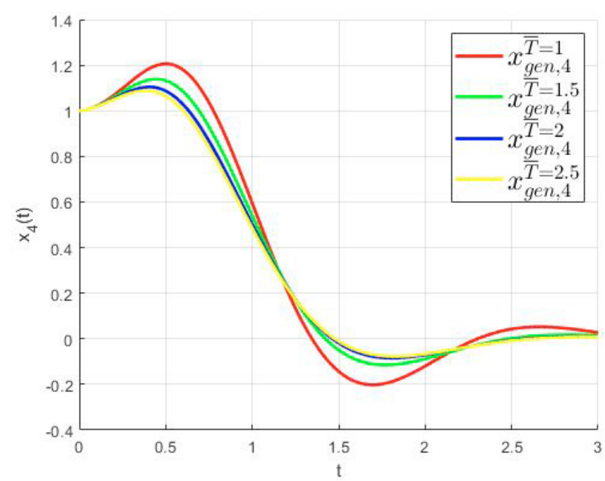

Fig. 3. $y(t)$ for the control problem with continuous updating with different lengths of the information horizon.

respectively. All data sets are generated based on the same cost function parameters (23) and system dynamics (20). We exemplarily show the corresponding optimal trajectories of the control in Fig. 2 as well as the lateral distance from the middle lane and the steering wheel angle in Fig. 3 and Fig. 4, respectively.

5.3 Inverse Optimal Control Problem

Now we solve the inverse optimal control problem (18) using the approach with continuous updating to obtain an estimate of θ as defined in (22). We use the data set D2, i.e. the optimal control trajectories with continuous updating with $\overline{T} = 1.5$, as observed trajectories $\hat{x}(t)$ and $\hat{u}(t)$ on the interval $[0, 10]$. In order to show the significance of identifying $\overline{T}$ asides from the cost function matrices, we present the optimal $[Q, R]$ for the different fixed values of $\overline{T}$ defined in (24). Furthermore, we show the corresponding values of the objective function Div (18). In order to avoid ambiguity of the parameters, we normalized all cost function parameters with respect to $R = [1.00]$. The results are given by

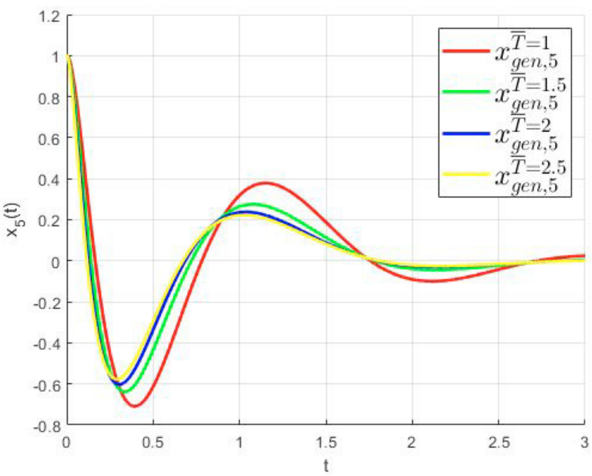

Fig. 4. $\delta(t)$ for the control problem with continuous updating with different lengths of the information horizon.

$$\begin{aligned}
Div^{\overline{T}=1} &= 7.7944 \cdot 10^{-9}, R = 1.00 \\
Q^{\overline{T}=1} &= diag(17.11, 38.56, 727.49, 91.64, 65.98, 2.50), \\
Div^{\overline{T}=1.5} &= 5.4986 \cdot 10^{-13}, R = 1.00 \\
Q^{\overline{T}=1.5} &= diag(11.95, 9.91, 1.00, 40.00, 10.01, 1.00), \\
Div^{\overline{T}=2} &= 0.017448, R = 1.00 \\
Q^{\overline{T}=2} &= diag(0.00, 0.00, 0.00, 18.54, 0.00, 0.34), \\
Div^{\overline{T}=2.5} &= 0.11415, R = 1.00 \\
Q^{\overline{T}=2.5} &= diag(0.00, 0.00, 0.00, 9.79, 0.00, 0.02).
\end{aligned} \tag{25}$$

The inverse optimal control approach introduced in Section 4, where we also identify the information horizon $\overline{T}$. The solution of (18) is calculated using the Sequential Quadratic Programming (SQP) method in the `fmincon` solver of MATLAB, where we constrained the parameters to be positive. The identification results are

$$\begin{aligned}
\overline{T}_{opt} &= 1.4815 \\
Div^{\overline{T}_{opt}} &= 1.3065 \cdot 10^{-10}, \; R = 1.00, \\
Q^{\overline{T}_{opt}} &= diag(177.12, 8.82, 9.84, 41.26, 10.99, 1.03).
\end{aligned} \tag{26}$$

Based on the identified parameters $Q^{\overline{T}_{opt}}$ and $\overline{T}_{opt}$, we generate the trajectories $u_{id}^{\overline{T}_{opt}}$ and $x_{id}^{\overline{T}_{opt}}$ corresponding to the optimal control with continuous updating. Fig. 5 shows the original optimal control (from data set D2, black line) and the optimal control with continuous updating from the identified parameters (yellow line). Similarly, in Fig. 6 and Fig. 7 the ground truth and identified model trajectories are shown for the lateral distance from the middle lane $y(t)$ and the steering wheel angle $\delta(t)$.

5.4 Discussion

From Fig. 2, it is noticeable that the optimal control trajectories are different depending on the assumed length of the information horizon $\overline{T}$. Additionally, Fig. 3 and Fig. 4 show that more information leads to a more effective stabilization of the states.

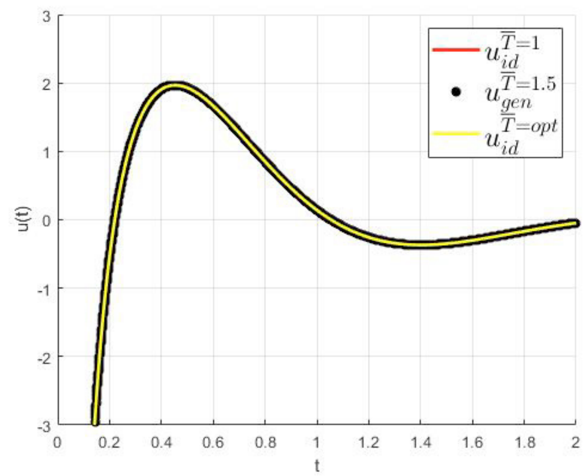

Fig. 5. Ground truth control (black line) and identified optimal control with continuous updating (yellow line) generated with $\overline{T}_{opt} = 1.4815$ and $Q^{\overline{T}_{opt}}$.

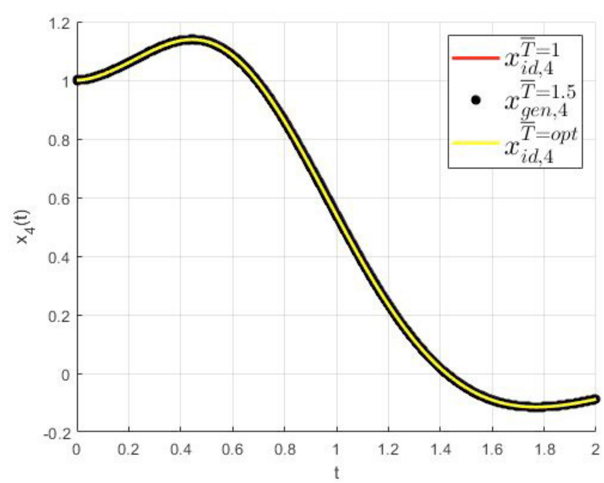

Fig. 6. Ground truth $y(t)$ (black line) and identified optimal control with continuous updating (yellow line) generated with $\overline{T}_{opt} = 1.4815$ and $Q^{\overline{T}_{opt}}$.

The solutions in (25) indicate that, if we assume that the human acts according to a moving information horizon, then the assumption of a fixed—possibly wrong —information horizon may lead to suboptimal results. This shows the importance of identification of optimal value of information horizon $\overline{T}$ from the observed data, which we were able to conduct by means of our performed inverse optimal control approach. The identified cost function parameters $Q^{\overline{T}_{opt}}$ with optimal $\overline{T}$ are the closest to the ground truth defined in (23), but differ especially in the first diagonal element of matrix Q, the parameter of the slip angle deviation. In this maneuver, the value of the slip angle does not affect the other states considerably. Therefore, it is a challenge to identify the corresponding parameters correctly. Nevertheless, the identified parameters were still able to explain the ground truth trajectories adequately.

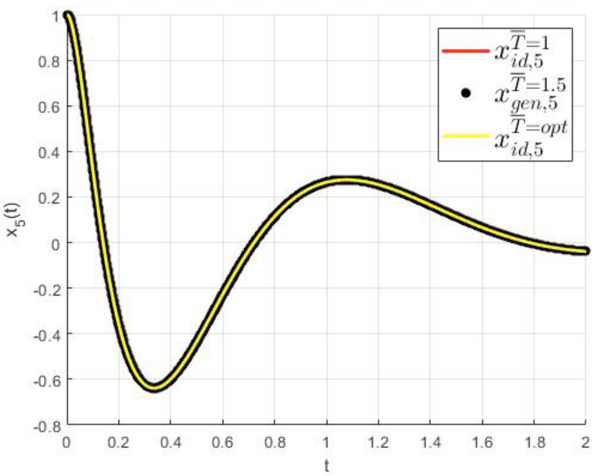

Fig. 7. Ground truth $\delta(t)$ (black line) and identified optimal control with continuous updating (yellow line) generated with $\overline{T}_{opt} = 1.4815$ and $Q^{\overline{T}_{opt}}$.

6. CONCLUSION

In this paper, we presented an optimal control approach and the corresponding inverse problem based on continuous updating, where the decision maker updates his behavior based on the new information available which arises from a shifting time horizon. We showed simulations results which show the applicability of the approach and also highlight the importance of identifying the real value of information horizon $\overline{T}_{opt}$. Our approach provides a means of more profound modeling of engineering systems with humans. Future work will focus on the test of inverse optimal control with continuous updating using real measured data from different human drivers.

ACKNOWLEDGMENTS

Research of Ovanes Petrosian and Ildus Kuchkarov was supported by a grant from the Russian Science Foundation (Project No 18-71-00081).

REFERENCES

Aghasadeghi, N. and Bretl, T. (2014). Inverse optimal control for differentially flat systems with application to locomotion modeling. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, 6018–6025. IEEE.

Bemporad, A., Morari, M., Dua, V., and Pistikopoulos, E. (2002). The explicit linear quadratic regulator for constrained systems. *Automatica*, 38(1), 3–20.

Cole, D.J., Pick, A.J., and Odhams, A.M.C. (2006). Predictive and linear quadratic methods for potential application to modelling driver steering control. *Vehicle System Dynamics*, 44(3), 259–284.

El-Hussieny, H., Abouelsoud, A., Assal, S.F., and Megahed, S.M. (2016). Adaptive learning of human motor behaviors: An evolving inverse optimal control approach. *Engineering Applications of Artificial Intelligence*, 50, 115–124.

El-Hussieny, H. and Ryu, J.H. (2018). Inverse discounted-based LQR algorithm for learning human movement behaviors. *Applied Intelligence*, 1489–1501.

Engwerda, J. (2005). *LQ Dynamic Optimization and Differential Games*. Willey, New York.

Flad, M., Otten, J., Schwab, S., and Hohmann, S. (2014). Steering driver assistance system: A systematic cooperative shared control design approach. In *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3585–3592.

Goodwin, G., Seron, M., and J.A., D. (2005). *Constrained Control and Estimation: An Optimisation Approach*. Springer, New York.

Gray, A., Gao, Y., Lin, T., Hedrick, J.K., and Borrelli, F. (2013). Stochastic predictive control for semi-autonomous vehicles with an uncertain driver model. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, 2329–2334.

Gromova, E. and O.L., P. (2016). Control of information horizon for cooperative differential game of pollution control. *2016 International Conference Stability and Oscillations of Nonlinear Control Systems (Pyatnitskiy's Conference)*.

Hempel, A., Goulart, P., and Lygeros, J. (2015). Inverse parametric optimization with an application to hybrid system control. *IEEE Transactions on Automatic Control*, 60(4), 1064–1069.

Inga, J., Eitel, M., Flad, M., and Hohmann, S. (2018). Evaluating Human Behavior in Manual and Shared Control via Inverse Optimization. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2699–2704.

Inga, J., Flad, M., Diehm, G., and Hohmann, S. (2015). Gray-Box Driver Modeling and Prediction: Benefits of Steering Primitives. In *2015 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 3054–3059.

Inga, J., Köpf, F., Flad, M., and Hohmann, S. (2017). Individual human behavior identification using an inverse reinforcement learning method. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 99–104.

Jean, F. and Maslovskaya, S. (2018). Inverse optimal control problem: the linear-quadratic case. In *2018 IEEE Conference on Decision and Control (CDC)*, 888–893. IEEE, FL, USA.

Johnson, M., Aghasadeghi, N., and Bretl, T. (2013). Inverse optimal control for deterministic continuous-time nonlinear systems. In *52nd IEEE Conference on Decision and Control*, 2906–2913.

Kuchkarov, I. and Petrosian, O. (2019). On class of linear quadratic non-cooperative differential games with continuous updating. *Lecture Notes in Computer Science*, 11548, 635–650.

Kwon, W., Bruckstein, A., and Kailath, T. (1982). Stabilizing state-feedback design via the moving horizon method. *21st IEEE Conference on Decision and Control*.

Kwon, W. and Pearson, A. (1977). A modified quadratic cost problem and feedback stabilization of a linear system. *IEEE Transactions on Automatic Control*, 22(5), 838–842.

Kwon, W. and Han, S. (2005). *Receding Horizon Control: Model Predictive Control for State Models*. Springer, New York.

Mayne, D. and Michalska, H. (1990). Receding horizon control of nonlinear systems. *IEEE Transactions on Automatic Control*, 35(7), 814–824. doi:10.1109/9.57020.

Mombaur, K., Truong, A., and Laumond, J.P. (2010). From human to humanoid locomotionan inverse optimal control approach. *Autonomous Robots*, 28(3), 369–383.

Pauwels, E., Henrion, D., and Lasserre, J.B.B. (2014). Inverse optimal control with polynomial optimization. *53rd IEEE Conference on Decision and Control*, 5581–5586.

Petrosian, O. and Gromova, E. (2018). Cooperative differential games with dynamic updating. *IFAC-PapersOnLine*, 51(32), 413–417.

Petrosian, O. and Kuchkarov, I. (2019). About the looking forward approach in cooperative differential games with transferable utility. *Frontiers of Dynamic Games: Game Theory and Management, St. Petersburg, 2018*, 175–208.

Petrosian, O., Shi, L., Li, Y., and Gao, H. (2019). Moving information horizon approach for dynamic game models. *Mathematics*, 7(12). doi:10.3390/math7121239.

Petrosian, O. and Tur, A. (2019). Hamilton-jacobi-bellman equations for non-cooperative differential games with continuous updating. In *Mathematical Optimization Theory and Operations Research*, 178–191.

Petrosian, O. (2016). Looking forward approach in cooperative differential games. *International Game Theory Review*, (18), 1–14.

Petrosian, O. and Barabanov, A. (2017). Looking forward approach in cooperative differential games with uncertain-stochastic dynamics. *Journal of Optimization Theory and Applications*, 172, 328–347.

Petrosian, O., Nastych, M., and Volf, D. (2018). Non-cooperative differential game model of oil market with looking forward approach. *Frontiers of Dynamic Games, Game Theory and Management, St. Petersburg, 2017, eds. L. A. Petrosyan, V. V. Mazalov, N. Zenkevich, Birkhuser, Basel, 2018.*

Priess, M.C., Conway, R., Jongeun Choi, Popovich, J.M., and Radcliffe, C. (2015). Solutions to the Inverse LQR Problem With Application to Biological Systems Analysis. *IEEE Transactions on Control Systems Technology*, 23(2), 770–777.

Ramadan, A., Choi, J., and Radcliffe, C.J. (2016). Inferring human subject motor control intent using inverse MPC. In *2016 American Control Conference (ACC)*, 5791–5796.

Rawlings, J. and Mayne, D. (2009). *Model Predictive Control: Theory and Design*. Nob Hill Publishing, Madison.

Shaw, L. (1979). Nonlinear control of linear multivariable systems via state-dependent feedback gains. *IEEE Transactions on Automatic Control*, 24(1), 108–112. Conference Name: IEEE Transactions on Automatic Control.

Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, 7(9), 907–915.

Wang, L. (2005). *Model Predictive Control System Design and Implementation Using MATLAB*. Springer, New York.

Yeung, D. and Petrosian, O. (2017). Cooperative stochastic differential games with information adaptation. *International Conference on Communication and Electronic Information Engineering.*

Zhang, H., Zhao, W., and Wang, J. (2019). Fault-Tolerant Control for Electric Vehicles With Independently Driven in-Wheel Motors Considering Individual Driver Steering Characteristics. *IEEE Transactions on Vehicular Technology*, 68(5), 4527–4536.