International Congress
Motor Vehicles & Motors 2024
Kragujevac, Serbia
October 10th - 11th, 2024

Motor Vehicles & Motors

MVM2024-NNN

Vojislav Filipovic[1]
Milan Matijevic[2]
Dragan Kostic[3]

DIGITAL PREVIEW CONTROLLER DESIGN USING REINFORCEMENT LEARNING

ABSTRACT: This paper discusses the design of a preview LQ controller. We assume a linear time-invariant continuous model of controlled mechatronics subsystem. The model is discretized using a generalized hold function (GHF), which is defined by its impulse response. This approach can improve the closed loop gain margin, among other benefits. The plant model is then expanded to include the preview part of the system, and the LQ controller design methodology is applied. An incremental control signal is introduced in the criterion, which adds integral action to the controller. The digital controller is obtained offline as a solution to the Riccati equation. This solution provides both the feedback and feedforward controllers. Finally, using adaptive dynamic programming, we convert the offline procedure into an online procedure, resulting in an intelligent controller. The analysis is supported by an illustrative example involving a general mechatronic subsystem. The results highlight the potential of designed digital controller to significantly enhance the performance of manufacturing systems. The obtained controller can be related to the ILC (Iterative Learning Control) family of control algorithms, which are typically used in manufacturing systems.

KEYWORDS: Manufacturing systems, Preview control, LQ, Reinforcement learning, Mechatronics systems

INTRODUCTION

Preview control is used to solve tracking or rejection problems under the assumption that the signals to be tracked or rejected are available a priori by a certain amount of time [2]. The first consideration of preview control in the form of examples is presented in [14]. Preview control for state space models was first proposed in [15]. In reference [13], a vehicle suspension problem is considered as a disturbance rejection problem. In [12,18], it is shown that the tracking problem in autonomous vehicles can be addressed as a preview control problem or as model predictive control using the preview-follower theory algorithm. An overview of digital tracking control in the field of preview control is presented in [16]. Applications include robot control [17], control of linear direct current motors [3], process control [9], and many other areas.

In this paper, we assume a known continuous-time model of the system. The implementation of the controller is digital, and therefore, the conversion of the continuous-time model to a discrete-time model is important. In this context, the form of the hold element is very important. In this paper, we consider the generalized hold function (GHF), which can be characterized by its impulse response when a unitary discrete-time impulse is used as input. The GHF is the subject of many papers and books [8, 6, 19, 20]. Among other benefits, the GHF can improve gain margin and robustness of the closed loop [19]. It is important to note that zero-order hold (ZOH) and first-order hold (FOH) are special cases of GHF.

The main topic of this paper is the design of a reinforcement learning LQ controller for the tracking problem. The first step is to design standard LQ controllers [1] for the tracking problem, where possible, using future information

[1] *Vojislav Filipovic, University of Kragujevac, Faculty of Mechanical and Civil Engineering in Kraljevo, Serbia*
[2] *Milan Matijevic, University of Kragujevac, Faculty of Engineering, Sestre Janjic 6, Kragujevac, Serbia, matijevic@kg.ac.rs*
[3] *Dragan Kostic, Enabling Technology Mechatronics, ASMPT Center of Competency, Netherlands*

(segment of the desired trajectory profile). This can improve tracking accuracy. Such problems are considered in references [4] and [9]. In this paper, we introduce the constraint that all closed-loop eigenvalues have magnitudes less $\lambda \in (0, 1]$) (specified degree of stability). The value of λ may be chosen by the designer [1]. This approach provides a fast response time for the system and greater tolerance to nonlinearity.

The main step in the design of the digital preview controller is to convert the tracking problem into a regulator problem. For that purpose, an augmented system is introduced (a model of the considered system extended with the model of the reference). The result is based on dynamic programming and consists of a two-degree-of-freedom regulator (feedback regulator and feedforward regulator). The observation is that the design of the regulator is based on the off-line solution of the algebraic Riccati equation.

The last part of the paper is devoted to converting the off-line solution of the Riccati equation to an on-line solution. This is accomplished with adaptive dynamic programming, which is the main topic in reinforcement learning [11, 7].

The main contribution of this paper is the design of an intelligent regulator based on the GHF and the concept of the degree of stability of the system.

DISCRETE-TIME MODEL OF CONTINUOUS SYSTEM

The sampling process is important for representing continuous-time systems with discrete-time models. Figure 1 shows a single-input single-output (SISO) system.

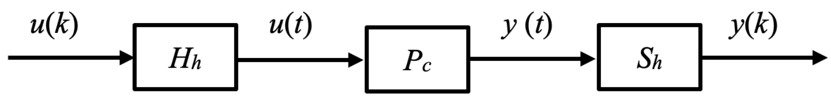

Figure 1 SISO discrete-time system

In Fig. 1 P_c represents a continuous linear time-invariant (LTI) plant, H_h represents hold element and S_h is the ideal sampler. Both H_h and S_h operate with the sampling period h. The mathematical description of discrete-time systems is

$$P = S_h P_c H_h \tag{1}$$

The hold element H_h is very important. In this paper, we consider the Generalized Hold Function (GHF). As pointed out in many references, the application of GHF has several benefits. According to [19], it can significantly increase the gain margin and robustness of the closed-loop system, and it makes a significant contribution to the field of multivariable systems [8]. In this paper, the Generalized Hold Function (GHF) is characterized by its impulse response. Fig. 2 illustrates the impulse response GHF during the first sampling time. It is shown that the impulse response during the first sampling time can be a nonlinear function of time. The Zero-Order Hold (ZOH) and First-Order Hold (FOH) are special cases of GHF.

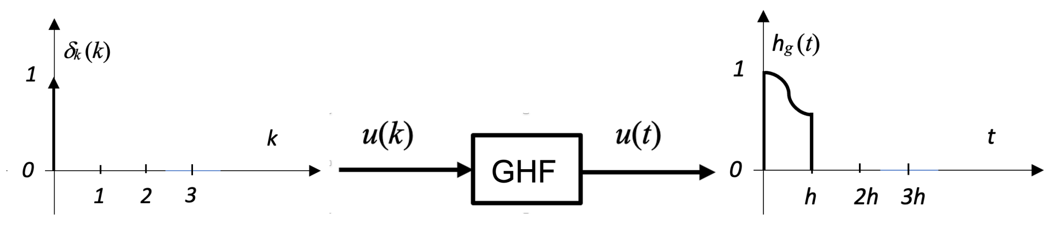

Figure 2 Principle of GHF

In the Fig. 2 input is

$$u(k) = \delta_k(k) = \begin{cases} 1 & k = 0 \\ 0 & k \neq 0 \end{cases} \tag{2}$$

and output of the GHF is the impulse response $\boldsymbol{h_g(t)}$ during the first sampling time. The continuous-time signal $\boldsymbol{u(t)}$ is

$$u(t) = \sum_{k=-\infty}^{\infty} h_g(t - kh) u(k) \tag{3}$$

Let us suppose that the continuous-time model of the plant P_c in state-space form for a SISO system is given

$$\dot{x}(t) = A_c x(t) + b_c u(t) \tag{4}$$

$$y(t) = c\,x(t)$$

If the GHF is used to generate the input $\boldsymbol{u}(\boldsymbol{t})$, the equivalent discrete-time model is given by [20]

$$\begin{gathered} x(k+1) = Ax(k) + bu(k) \\ y(k) = c\,x(k) \end{gathered} \tag{5}$$

where

$$A = e^{A_c h}, b = \int_0^h e^{A_c(h-\tau)} b_c h_g(\tau) d\tau \tag{6}$$

Remark 1. An example of a GHF is a piecewise impulse response

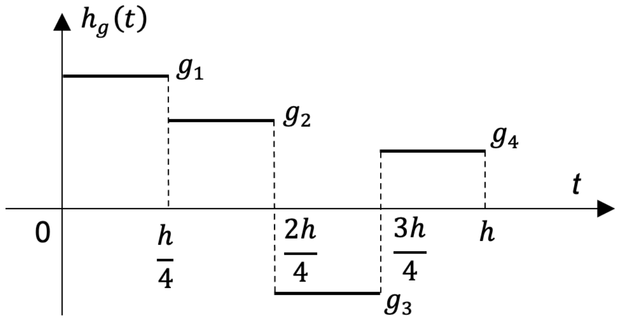

Figure 3 Impulse response of a picewise GHF for *N*=4 during one sample time

Designing the GHF as an *N*-piecewise GHF (illustrated in Fig. 3) affects the form of the discrete model (6) of the controlled plant P_c in the following way

$$h_g(t) = \begin{cases} g_1 & 0 \le t < \dfrac{h}{N} \\ \vdots & \\ g_N & \dfrac{(N-1)h}{N} \le t < h \end{cases} \quad => \quad b = \sum_{l=1}^{N} g_l \int_{\frac{(l-1)h}{N}}^{\frac{lh}{N}} e^{A_c(h-\tau)} b_c d\tau \tag{7}$$

DIGITAL REVIEW REGULATOR

In this control strategy, future information about the reference trajectory is incorporated into the controller design to improve tracking accuracy. This is combined with a linear quadratic (LQ) regulator. For this case, an appropriate performance index is used, and an augmented state-space model includes the available future demands as part of the state vector. We assume that, at each time, the reference trajectory includes N_p future values $r(k+1), \dots, r(k+N_p)$.

We now introduce the following assumption:

Assumption: Future values of the reference trajectory beyond time $k+N_p$ are approximated by $r(k+i) = r(k+N_p)$ for $i = N_p+1, N_p+2, \dots$.

For system (5), we use the following criterion:

$$J = \sum_{k=0}^{\infty} \lambda^{-2k} (q_e e^2(k) + \Delta x^{\mathrm{T}}(k) Q_x \Delta x(k) + r_u \Delta u^2(k)) \tag{8}$$

where $q_0 \ge 0$, $r_u > 0$, $\Delta x(k) = x(k) - x(k-1)$, $\Delta u(k) = u(k) - u(k-1)$, and λ represents degree of stability of closed loop system. Now we will transform system (5)

$$\hat{x}(k) = \lambda^{-k} x(k), \;\; \hat{u}(k) = \lambda^{-k} u(k) \tag{9}$$

For system (5), we have

$$\lambda^{-(k+1)} x(k+1) = \lambda^{-1} A\big(\lambda^{-k} x(k)\big) + \lambda^{-1} b\big(\lambda^{-k} u(k)\big) = \lambda^{-1} A\hat{x}(k) + \lambda^{-1} b\hat{u}(k) \tag{10}$$

The transformed system has the form

$$\hat{x}(k+1) = \lambda^{-1}A\hat{x}(k) + \lambda^{-1}b\hat{u}(k)$$
$$\hat{y}(k) = c\,\hat{x}(k) \tag{11}$$

After system transformation, criterion (8) takes the form

$$J = \sum_{k=0}^{\infty}\left(q_e\hat{e}^2(k) + \Delta\hat{x}^{\mathrm{T}}(k)Q_x\Delta\hat{x}(k) + r_u\Delta\hat{u}^2(k)\right) \tag{12}$$

Criterion (12) includes the incremental form of the state vector. As known, this form of criterion provides an integral term in the regulator. Now we find the incremental form of system (11). Let us note that

$$\Delta\hat{x}(k+1) = \hat{x}(k+1) - \hat{x}(k),\ \ \Delta\hat{u}(k+1) = \hat{u}(k+1) - \hat{u}(k) \tag{13}$$

From equations (11) and (13), it follows that

$$\begin{aligned}\Delta\hat{x}(k+1) &= \hat{x}(k+1) - \hat{x}(k) = \lambda^{-1}A\hat{x}(k) + \lambda^{-1}b\hat{u}(k) - \lambda^{-1}A\hat{x}(k-1) - \lambda^{-1}b\hat{u}(k-1)\\ &= \lambda^{-1}A\left(\hat{x}(k) - \hat{x}(k-1)\right) + \lambda^{-1}b\left(\hat{u}(k) - \hat{u}(k-1)\right) = \lambda^{-1}A\Delta\hat{x}(k) + \lambda^{-1}b\Delta\hat{u}(k)\end{aligned} \tag{14}$$

The incremental form of system (11) is

$$\begin{aligned}\Delta\hat{x}(k+1) &= \lambda^{-1}A\,\Delta\hat{x}(k) + \lambda^{-1}b\,\Delta\hat{u}(k)\\ \Delta\mathrm{y}(k) &= c\,\Delta\hat{x}(k)\end{aligned} \tag{15}$$

Remarks 2. The incremental system description equation (15) allows us to consider the following system:

$$\begin{aligned}x(k+1) &= Ax(k) + bu(k) + w(k)\\ y(k) &= c\,x(k)\end{aligned} \tag{16}$$

where $w(k)$ is a constant disturbance. The error signal is

$$\begin{aligned}\Delta\hat{e}(k+1) &= \Delta r(k+1) - \Delta y(k+1) = \Delta r(k+1) - c\,\Delta\hat{x}(k+1)\\ &= \Delta r(k+1) - c\left(\lambda^{-1}A\,\Delta\hat{x}(k) + \lambda^{-1}b\,\Delta\hat{u}(k)\right)\\ &= \Delta r(k+1) - \lambda^{-1}cA\,\Delta\hat{x}(k) - \lambda^{-1}cb\,\Delta\hat{u}(k)\end{aligned} \tag{17}$$

From the last equation, it follows that

$$\hat{e}(k+1) = \hat{e}(k) + \Delta r(k+1) - \lambda^{-1}cA\,\Delta\hat{x}(k) - \lambda^{-1}cb\,\Delta\hat{u}(k) \tag{18}$$

Using relations (15) and (18), it follows that

$$\begin{bmatrix}\hat{e}(k+1)\\ \Delta\hat{x}(k+1)\end{bmatrix} = \begin{bmatrix}1 & -\lambda^{-1}cA\\ 0 & \lambda^{-1}A\end{bmatrix}\begin{bmatrix}\hat{e}(k)\\ \Delta\hat{x}(k)\end{bmatrix} + \begin{bmatrix}-\lambda^{-1}cb\\ \lambda^{-1}b\end{bmatrix}\Delta\hat{u}(k) + \begin{bmatrix}0\\1\end{bmatrix}\Delta r(k+1) \tag{19}$$

Now we will design a command generator system that models the preview part of the system and has the following form:

$$\begin{aligned}x_r(k+1) &= A_r x_r(k)\\ y_r(k) &= c_r x_r(k)\end{aligned} \tag{20}$$

where

$$A_r = \begin{bmatrix}0 & 1 & 0 & \dots & 0\\ 0 & 0 & 1 & \dots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & 1\\ 0 & 0 & 0 & \cdots & 0\end{bmatrix}$$

The matrix A_r implements a shift register operation. The state of the command generator $x_r(k)$ is composed of the sampled values of the reference signal over the preview horizon of length N_p. The form of the vector $x_r(k)$ is

$$x_r(k) = \begin{bmatrix}\Delta r(k+1)\\ \Delta r(k+2)\\ \vdots\\ \Delta r\left(k+N_p\right)\end{bmatrix},\ \ \hat{x}_r(k+1) = \lambda^{-1}A_r\hat{x}_r(k),\ \ \hat{x}_r(k) = \lambda^{-k}x_r(k) \tag{21}$$

Now define the augmented state vector

$$\hat{x}_1(k) = [\hat{e}(k)\quad \Delta\hat{x}^{\mathrm{T}}(k)\quad \hat{x}_r^{\mathrm{T}}(k)]^{\mathrm{T}},\ \ \ x_1(k) = [e(k)\quad \Delta x^{\mathrm{T}}(k)\quad x_r^{\mathrm{T}}(k)]^{\mathrm{T}} \tag{22}$$

From equations (8) and (20), it follows that

$$\hat{x}_1(k+1) = \left[\begin{array}{cc|ccc} 1 & -\lambda^{-1}cA & 1 & 0 \; \ldots & 0 \\ 0 & \lambda^{-1}A & 0 & 0 \; \ldots & 0 \\ \hline \multicolumn{2}{c|}{0} & \multicolumn{3}{c}{\lambda^{-1}A_r} \end{array}\right] \hat{x}_1(k) + \left[\begin{array}{c} -\lambda^{-1}cb \\ \lambda^{-1}b \\ \hline 0 \end{array}\right] \Delta\hat{u}(k) = A_1\hat{x}_1(k) + b_1\Delta\hat{u}(k) \tag{23}$$

The performance index (12) now takes the following form:

$$J_r = \sum_{k=0}^{\infty} \left(\hat{x}_1^{\mathrm{T}}(k) Q_1 \hat{x}_1(k) + r_u \Delta\hat{u}^2(k) \right) \tag{24}$$

where

$$Q_1 = \begin{bmatrix} q_e & \square & 0 \\ \square & Q_x & \square \\ 0 & \square & 0 \end{bmatrix}$$

The optimization problem is standard, involving the minimization of criterion (24) subject to constraint (23). According to [13] and relations (23) and (24), the optimal regulator has the following structure:

$$\Delta\hat{u}^0(k) = -(r_u + b_1^{\mathrm{T}} P_1 b_1)^{-1} b_1^{\mathrm{T}} P_1 b_1 \hat{x}_1(k) \tag{25}$$

where P_1 is a symmetric matrix that is the solution to the following algebraic matrix Riccati equation:

$$P_1 = A_1^{\mathrm{T}} P_1 A_1 - A_1^{\mathrm{T}} P_1 b_1 (r_u + b_1^{\mathrm{T}} P_1 b_1)^{-1} b_1^{\mathrm{T}} P_1 A_1 + Q_1 \tag{26}$$

From relation (26) and the structure of the regulation we have

$$\Delta\hat{u}^0(k) = -[k_e \quad k_x \quad k_r]\hat{x}_1(k) \tag{27}$$

Now we explicitly find gains *K* of the regulator. To do this, we define

$$b_2 = \begin{bmatrix} -\lambda^{-1}cb \\ \lambda^{-1}b \end{bmatrix}, F_2 = \begin{bmatrix} \lambda^{-1}cA \\ \lambda^{-1}A \end{bmatrix}, Q_2 = \begin{bmatrix} q_e & 0 \\ 0 & Q_x \end{bmatrix}, I_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, A_2 = [I_2 \quad F_2] \tag{28}$$

Now we will adapt the results from [16] to our case. The optimal incremental control is given by

$$\Delta\hat{u}^0(k) = -k_i\hat{e}(k) - k_x\Delta x(k) - \lambda^{-k}\sum_{l=1}^{N_p}\left(k_r(l)\Delta r(k+l)\right)$$

From which it follows that

$$\Delta u^0(k) = -k_i e(k) - k_x \Delta x(k) - \sum_{l=1}^{N_p} k_r(l)\Delta r(k+l) \tag{29}$$

$$k_i = -(r_u + b_2^{\mathrm{T}} P_2 b_2)^{-1} b_2^{\mathrm{T}} P_2 I_2 \tag{30}$$

$$k_x = -(r_u + b_2^{\mathrm{T}} P_2 b_2)^{-1} b_2^{\mathrm{T}} P_2 F_2 \tag{31}$$

$$k_r(l) = \begin{cases} k_r(1) = -k_i \\ \begin{cases} k_r(l) = (r_u + b_2^{\mathrm{T}} P_2 b_2)^{-1} b_2^{\mathrm{T}} x(l-1), \quad l = 2, \ldots, N_p \\ x(l) = A_{2c} x(l-1), \;\; l = 2, \ldots, N_p \\ x(1) = -A_{2c} P_2 I_2 \\ A_{20} = A_2 - b_2 (r_u + b_2^{\mathrm{T}} P_2 b_2)^{-1} b_2^{\mathrm{T}} P_2 A_2 \end{cases} \end{cases} \tag{32}$$

where the matrix P_2 is the non-negative definite solution of the following algebraic Riccati equation

$$P_2 = A_2^{\mathrm{T}} P_2 A_2 - A_2^{\mathrm{T}} P_2 b_2 (r_u + b_2^{\mathrm{T}} P_2 b_2)^{-1} b_2^{\mathrm{T}} P_2 A_2 + Q_2 \tag{33}$$

From equation (29), we will now find the position (or whole value) form of the optimal control. Let us note that

$$e(k)=\sum_{i=0}^{k}e(i)-\sum_{i=0}^{k-1}e(i) \tag{34}$$

Algorithm (29) takes the following form:

$$\begin{aligned} u^0(k)-u^0(k-1) \\ &= -k_i\left(\sum_{i=0}^{k}e(i)-\sum_{i=0}^{k-1}e(i)\right)-k_x\big(x(k)-x(k-1)\big) \\ &-\sum_{l=1}^{N_p}k_r(l)(r(k+l)-r(k+l-1)) = \left\{-k_i\sum_{i=0}^{k}e(i)-k_xx(k)-\sum_{l=1}^{N_p}k_r(l)r(k+l)\right\} \\ &+\left\{k_i\sum_{i=0}^{k-1}e(i)+k_xx(k-1)+\sum_{l=1}^{N_p}k_r(l)r(k+l-1)\right\} \end{aligned} \tag{35}$$

From the last relation, one can derive the position format algorithm:

$$u^0(k)=-k_i\sum_{i=0}^{k}e(i)-k_xx(k)-\sum_{l=1}^{N_p}k_r(l)\Delta r(k+l) \tag{36}$$

Now we will present the algorithm for designing the optimal preview regulator.

Algorithm for optimal preview regulator

1. Given the continuous-time system (4)
2. Design of discrete-time system (5) using the GHF (with relations (4)-(7)) for chosen sampling time *h*.
3. Choose parameters $q_e \geq 0, r_u > 0, Q_x \geq 0, \lambda\epsilon(0,1]$.
4. Solve the Riccati equation (33) to determine the matrix P_2
5. Determine the regulator's gains k_i, k_x, k_r using relations (30) – (32).

Remark 3. It is possible to establish a connection between optimal preview and iterative learning control (ILC) as showed in [21].

Remark 4. The relation between the considered optimal preview regulator and predictive regulators is presented in [5].

MODEL-BASED INTELLIGENT PREVIEW REGULATOR

In the previous section, the design of the digital preview regulator was based on the off-line solution of the Riccati equation. Here, we convert the off-line solution with on-line solution of the Riccati equation, as discussed in [10]. This methodology falls under reinforcement learning, which suggests generalized policy iteration [10].

We will now reconsider the system:

$$\hat{x}_1(k+1)=A_1\hat{x}_1(k)+b_1\Delta\hat{u}(k) \tag{37}$$

$$J_r(\hat{x}_1(k))=\sum_{k=0}^{\infty}\left(\hat{x}_1^{\mathrm{T}}(k)Q_1\hat{x}_1(k)+r_u\Delta\hat{u}^2(k)\right) \tag{38}$$

The objective to find the minimum of $J(\cdot)$. It is a well-known fact that the optimal value for the LQ regulator [1] is

$$\min_{\hat{x}_1(k)} J_r(\hat{x}_1(k))=\hat{x}_1^{\mathrm{T}}(k)P_3\hat{x}_1(k) \tag{39}$$

for some symmetric matrix P_3, which is determined by the solution of the algebraic Riccati equation.

Let us introduce the Bellman equation for the LQ regulator [10], using the facts that $\hat{u}(k)=\lambda^{-k}u(k), \hat{x}_1(k)=\lambda^{-k}x(k)$)

$$\hat{x}_1^{\mathrm{T}}(k)P_3\hat{x}_1(k)=\hat{x}_1^{\mathrm{T}}(k)Q_1\hat{x}_1(k)+r_u\Delta\hat{u}^2(k)+\hat{x}_1^{\mathrm{T}}(k+1)P_3\hat{x}_1(k+1) \tag{40}$$

The feedback is defined as

$$\Delta\hat{u}(k) = -k_u\hat{x}_1(k), \quad k_u = [k_e \quad k_x \quad k_r]^{\mathrm{T}} \tag{41}$$

We further have

$$\begin{aligned}\hat{x}_1^{\mathrm{T}}(k+1)P_3\hat{x}_1(k+1) &= \left(A_1\hat{x}_1(k) - b_1k_u\hat{x}_1(k)\right)^{\mathrm{T}}\left(A_1\hat{x}_1(k) - b_1k_u\hat{x}_1(k)\right) = \\ &= \hat{x}_1^{\mathrm{T}}(k)(A_1 - b_1k_u)^{\mathrm{T}}P_3(A_1 - b_1k_u)\hat{x}_1(k)\end{aligned} \tag{42}$$

$$\hat{x}_1^{\mathrm{T}}(k)Q_1\hat{x}_1(k) + r_u\Delta\hat{u}^2(k) = \hat{x}_1^{\mathrm{T}}(k)Q_1\hat{x}_1(k) + r_u\left(-k_u\hat{x}_1(k)\right)^{\mathrm{T}}\left(-k_u\hat{x}_1(k)\right) = \hat{x}_1(k)(Q_1 + r_uk_u^Tk_u)\hat{x}_1(k) \tag{43}$$

Based on equations (40) – (43), it follows that

$$(A_1 - b_1k_u)^{\mathrm{T}}P_3(A_1 - b_1k_u) - P_3 + Q_1 + r_uk_u^Tk_u = 0 \tag{44}$$

Let us write the Bellman equation in the following form

$$\hat{x}_1^{\mathrm{T}}(k)P_3\hat{x}_1(k) = \hat{x}_1^{\mathrm{T}}(k)Q_1\hat{x}_1(k) + r_u\Delta\hat{u}^2(k) + \left(A_1\hat{x}_1(k) + b_1\Delta\hat{u}(k)\right)^{\mathrm{T}}P_3\left(A_1\hat{x}_1(k) + b_1\Delta\hat{u}(k)\right) \tag{45}$$

If we differentiate of relation (45) with respect to $\Delta\hat{u}(k)$, we obtain

$$\Delta\hat{u}(k) + b_1^{\mathrm{T}}P_3\left(A_1\hat{x}_1(k) + b_1\Delta\hat{u}(k)\right) = 0$$

from which it follows that

$$\Delta\hat{u}^0(k) = -(r_u + b_1^{\mathrm{T}}P_3b_1)^{-1}b_1^{\mathrm{T}}P_3A_1\hat{x}_1(k) \tag{46}$$

We will now present the algorithm for the intelligent regulator using relations (44) and (46).

Algorithm for intelligent regulator

1. Given discrete-time system (23) with sampling time *h*
2. Choose $q_e \geq 0, r_u > 0, Q_x \geq 0, \lambda\epsilon(0,1], P_0 = I, \varepsilon > 0, j = 0, i = 1, 2, \dots, n-1; n > 0, n\epsilon N$; as well as initial regulator k_0 (not necessarily stabilizing)
3. Set:
   $(P_3)_j^0 = (P_3)_j$;
   $(P_3)_{j+1} = (P_3)_j^n$;
4. Solve the matrix equation
   $(P_3)_j^{i+1} = \left(A_1 - b_1k_{u,j}\right)^{\mathrm{T}}(P_3)_j^i\left(A_1 - b_1k_{u,j}\right) + Q_1 + r_uk_{u,j}^Tk_{u,j}$
5. Calculate the regulator gain
   $k_{u,j+1} = \left(r_u + b_1^{\mathrm{T}}(P_3)_{j+1}b_1\right)^{-1}b_1^{\mathrm{T}}(P_3)_{j+1}A_1$
6. Apply incremental control
   $\left(\Delta u^0(k)\right)_{j+1} = -k_{u,j+1}x_1(k)$
7. Stop if $\left\|k_{u,j+1} - k_{u,j}\right\| < \varepsilon$
   Otherwise, set *j=j+1* and return to step 3.

CONCLUSIONS

In this paper, we address the tracking problem under the assumption that the signal to be tracked is known a priori over a bounded interval of time. We first derive the discrete-time model from the continuous-time model using the Generalized Hold Function (GHF), determine the degree of stability, and incorporate the incremental structure of the model and preview formulation. This converts the tracking problem into a regulation problem.

Second, we solve the LQ regulator problem for the given model, using explicit expressions for the regulator gains ($k_e \quad k_x \quad k_r$). We consider two forms of algorithms: incremental and position, that have the structure of a PI regulator.

Third, we design an intelligent regulator based on reinforcement learning for the same model. Further investigations will focus on multivariable systems and continuous time systems.

REFERENCES

[1] Anderson B., Moore J.: “Optimal control. Linear quadratic methods”, Prentice-Hall, New Jersey, 1989.

[2] Birla N., Swarup A.: “Optimal preview control: a review”, Optimal Control Applications and Methods, Vol. 36, No. 2, 2015, 241-268.

[3] Egami T., Tsuchiya T.: “Disturbance suppression control with preview action of linear DC brushless motor”, IEEE Trans. on Industrial Electronics, Vol. 42, No. 5, 1995, 494-500.

[4] Farook A., Lmebeer D.: “Path following of optimal trajectories using preview control”, Proceedings of the 44th CDC and European Control Conference, Sevile, Spain, 2005, 2787-2992.

[5] Fang X., Chen W.H.: ”Model predictive control with preview recursive feasibility and stability”, IEEE Control Systems Letters, No.6, 2022, 2647-2052.

[6] Fener A., Goodwin G.C.: ”Generalized sample hold functions frequency domain analysis of robustness, sensitivity and intersample difficulties”, IEEE Transactions on Automatic Control, Vol. 39, No. 5, 1994, 1042-1047.

[7] Hua C.: “Reinforcement learning aided performance optimization of feedback control systems”, Springer, 2021.

[8] Kabamba P.T.: “Control of linear systems using generalized sampled data hold functions”, IEEE Trans. on Automatic Control, Vol. 32, 1987, 772-783.

[9] Katayama T., Omri T., Inove T., Kato T.: ”Design of an optimal controller for a discrete-time systems subject to previewable demand”, Int.J.Control, Vol. 41, No. 30, 1985, 677-699.

[10] Lewis F., Vrabie D.: “Reinforcement learning and adaptive dynamics programming for feedback control”, IEEE Circuits and Systems Magazine, Vol. 8, No. 3, 2009, 32-50.

[11] Li J., Lewis F., Fan J.: ”Reinforcement learning. Optimal feedback control with industrial applications”, Springer, 2023.

[12] Ravat A.K., Dhwan A., Tiwari M.: “Preview control for discrete-time systems”, in: Dhwan A. (eds), “Recent trends in electronics and communication”, Springer 2021, 1157-1166.

[13] Roh S.H., Park Y.: ”Stohastic optimal preview control of an active vehicle suspension”, Journal of Sound and Vibration, Vol. 220, No. 2. 1999. 313-330.

[14] Sheridan T.B.: “Three models of preview control”, IEEE Trans. On Human Factors in Electronics, Vol. 7, No. 2, 1966, 91-102.

[15] Tomizuka M., Whitney D.E.: “Optimal discrete finite preview problems (why and how is future information important?)”, ASME Journal of Dynamics Systems, Measurement and Control, 1975, 319-325.

[16] Tomizuka M.: “On the design of digital tracking controllers”, ASME Journal of Dynamic Systems, Measurement and Control, Vol, 115, No. 6, 1993, 412-418.

[17] Tsuchiya T., Shianda K., Abe S., Egami T.: ”Robot trajectory control and adaptive control”, Proceedings of the IEEE Workshop on Intelligent Robots, Tokyo, Japan, 1988: 125-130.

[18] Xu Y., Tang W., Chen B., Qui L., Yang R.: “A model predictive control with preview-follower theory algorithm for trajectory tracking control in autonomous vehicles”, Symmetry, Vol.13, 2021, 381-397.

[19] Yan W.Y., Anderson B.D.O, Bitmead R.R.: “On the gain margin improvement using dynamic compensation based on generalized sampled-data hold functions”, IEEE Transactions on Automatic Control, Vol. 39, No. 11, 1994, 2347-2354.

[20] Yuz J., Goodwin G.: “Sampled-data models for linear and nonlinear systems”, Springer, 2014.

[21] Zundert J., Oomen T.: “On inversion-based approaches for feedforward and ILC”, Mechatronics, Vol. 50, 2018, 282-291.