

Automatic track guidance of industrial trucks using AI-based controllers with disturbance compensation

Timm Sauer^{1†}, Manuel Gorks¹, Luca Spielmann¹, Klaus Zindler¹ and Ulrich Jumar²

¹Department of Engineering, University of applied sciences, Aschaffenburg, Germany

(E-mail: Timmanton.sauer@th-ab.de, Manuel.Gorks@th-ab.de, Luca.Spielmann@th-ab.de, Klaus.Zindler@th-ab.de)

²ifak - Institute of Automation and Communication, Magdeburg, Germany

(E-mail: Ulrich.Jumar@ifak.eu)

Abstract: This paper presents a new control concept for automatic track guidance of industrial trucks in intralogistic systems. It is based on Reinforcement Learning (RL), a method of Artificial Intelligence (AI). The presented approach is able to adapt itself to different industrial truck variants and to the associated specific vehicle parameters. In order to avoid starting the whole training of the controller for each truck variant from scratch, the training process is divided into two steps. In the first step, the controller is trained on a model using parameters of a nominal vehicle variant. Based on this, the control parameters are only fine-tuned in the second step. In this way the controller is adapted to the actual truck variant and the corresponding parameter values. Moreover, the influence of the disturbance variable of the system (path curvature) is compensated by considering this a priori knowledge within the control design. Therefore, the Artificial Neural Networks (ANN) of the RL controller and the observation vector are suitably adjusted. In this way, the occurring path curvatures can be considered in both training steps and the control parameters can be optimized accordingly. Thus, the influence of the disturbance variable can be compensated, which significantly improves the control quality. In order to demonstrate this, the new approach is compared to a RL control concept, not considering the disturbance variable and to a classical two-degrees-of-freedom (2DoF) control approach.

Keywords: Artificial intelligence, Automatic control, Intelligent transportation systems

1. INTRODUCTION

1.1. Problem description and requirements

In times of global economic markets and increasing competition, the automation of logistic processes is a basic requirement for corporate success. An important object of research and development is to increase the internal material flow via an autonomous and intelligent networked fleet, that usually consists of a wide variety of different individual truck variants. An essential element in this environment is the automatic track guidance of industrial trucks.

The principle of automatic steering control of a forklift is demonstrated in figure 1. First of all, the desired vehicle trajectory (predefined path) is calculated and stored as a data set on the real-time computer. The data record includes the necessary setpoint information for the automated vehicle guidance, such as the Cartesian Coordinates and curvature of the trajectory. The main objective is to guide the truck as accurately as possible along the path that only small lateral deviations occur. In [1] it is shown, that it is of benefit to the controllability of the system if a preview point P_p is guided instead of the vehicles center of gravity (CoG). For this purpose, P_p is defined in the preview distance l_p in front of the industrial truck's CoG [2]. The lateral deviation a_p corresponds to the distance between the preview point P_p and the reference point R_p on the predefined path.

Using this information, the controller calculates an appropriate control signal for the steering actuator in order to reduce the lateral deviation. The varying path curva-

ture during operation has a significant influence on the resulting lateral deviation of the industrial truck and thus on the automatic track guidance system. Since the path is defined in advance, this information represents a priori knowledge and should definitely be exploited for automatic track guidance. Hence, it should be taken into account by the control concept.

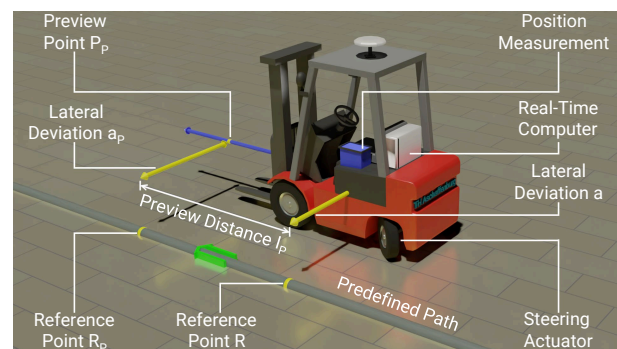


Fig. 1 Principle of automatic track guidance

The classical control design is based on a mathematical model that describes the dynamics of the controlled system (plant) as accurately as possible. In case of a heterogeneous fleet, a suitable model has to be derived for each vehicle variant. Based on this model, the control design has to be carried out for every single truck variant, which proves to be time-consuming. Consequently, a control concept for automatic track guidance of industrial trucks has to be developed that independently adapts to different industrial truck variants and moreover considers a priori knowledge, like the influence of the varying path curvature during operation.

† Timm Sauer is the presenter of this paper.

1.2. Related research

The papers [3] - [6] deal with the topic of automatic track guidance of industrial trucks, but each of them is focusing only on a single truck variant.

In the publications [7] - [10] a 2DoF control concept for automatic track guidance of vehicles is presented, that specifically considers the influence of the disturbance variable (path curvature) as a priori knowledge. These control structures consist of a linear disturbance compensation (feedforward controller FFC) combined with different kinds of feedback controllers (FBC). These approaches proved to be very effective, since the influence of the changing path curvature can almost be compensated. Compared to a closed loop control concept, significant advantages can be achieved using the 2DoF approach. Since both parts of the lateral controller (FFC and FBC) depend on the controlled system (section 3), this concept is suitable for only one single truck variant. In order to consider multiple forklift variants, new methods based on AI are used in addition to the classical adaptive control concepts given in [12] - [14]. An overview as well as a classification of the different AI approaches is given in [1]. The well-known RL control methods suffer from the fact, that a priori knowledge is not integrated in the training process [15], [16]. Therefore, a new approach has been presented in [1] that will be called Reinforcement Learning Control Considering a priori Plant Knowledge (RLCCPK) in the following. It's basic idea consists of integrating a priori knowledge of the controlled system into the training process. For this purpose, the training is divided into two steps. In the first step the controller is pre-trained on basis of a nominal model representing a priori knowledge of lateral dynamic vehicle behaviour. Since this model is derived for an industrial truck with average vehicle parameter values, in the second step a fine tuning of the control parameters is performed in order to adapt to the actual vehicle variant. In this way the efficiency of the whole training process is significantly increased. However, the RLCCPK approach considers a priori knowledge of the controlled system but neglects the influence of the varying path curvature during operation. Since the path is available in advance, this a priori knowledge should absolutely be taken into account by the control concept.

1.3. Main contribution and outline of this paper

This paper presents a new control concept for the automatic track guidance of industrial trucks which is based on RL. It adapts itself to different vehicle variants and also takes into account a priori knowledge. RL is implemented in form of the so-called Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm, as it proves to be suitable for the application of automatic track guidance [17]. The method of integrating a priori plant knowledge into the training process known from RLCCPK is extended to compensate the influence of the disturbance variable (known in advance) in analogy to a 2DoF control concept (subsection 1.2).

By means of an appropriate extension of the so-called observation vector (section 4) the path curvature is provided to the RL controller. Furthermore, the structures of the RL controller's ANN have to be adjusted in order to process the information of the observation vector. Thus, the control quality can significantly be improved. To demonstrate this, the control concept proposed in this paper, called Reinforcement Learning Control with Disturbance Compensation (RLCDC), is compared to the RLCCPK and a 2DoF control approach.

This paper is organized as follows. Section 2 introduces the used control structures as well as the modeling of the plant. In section 3 the design of the 2DoF controller is carried out using the root locus method. The fundamentals of RL as well as the AI-based control approaches (RLCCPK and RLCDC) will be introduced in section 4. Subsequently, the simulation results of the used control concepts are assessed (section 5). At the end of the paper, in section 6, the main conclusions are discussed.

2. CONTROL STRUCTURES AND MODELING OF THE PLANT

This section presents the structures of the used control concepts and the modeling of the plant. First of all, the fundamental structure of the lateral dynamic vehicle guidance system is presented. Based on this, the structure of the plant model is explained step by step. The following subsection 2.2 is dedicated to the structure of the 2DoF control concept, since it is used as a comparison control approach in this paper. Finally, the proposed AI-based control concept for the specific consideration of the disturbance variable is presented in analogy to the 2DoF approach in subsection 2.3.

2.1. Fundamental control structure and modeling of the plant

Figure 2 provides the fundamental control structure of the vehicle guidance system. The output of the lateral controller δ_{set} which is the first input signal of the controlled system is calculated with respect to lateral deviation a_p . As shown in [1], the preview concept is used in order to eliminate the non-minimum phase system behaviour, caused by the rear axle steering of the industrial truck. Therefore, the lateral deviation corresponds to the distance between the reference point R_p and the preview point P_p (figure 1). Obviously, it is necessary to calculate the position of R_p , which is done by means of the algorithms given in [7]. The curvature χ_p of the predefined path in the reference point R_p represents the second input of the plant and is considered as disturbance variable.

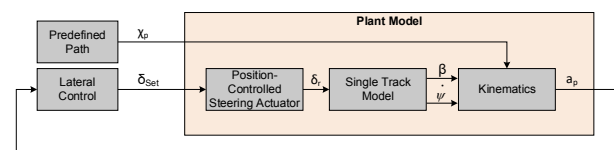


Fig. 2 Fundamental control structure

The plant model itself consists of three parts starting with the position controlled steering actuator which gets the calculated setpoints δ_{set} as its input signal and accordingly adjusts the rear axle steering angle δ_r . The second part of the plant is the so-called single track model that describes the lateral dynamic behaviour of the industrial truck depending on the steering angle δ_r [7] - [10], [19]. The last part represents the kinematics of the vehicle and describes the relative motion of the industrial truck with respect to the predefined path [7] - [10], [18]. The resulting lateral deviation a_p (output signal of the plant) with respect to the preview point P_p forms the input signal of the lateral controller.

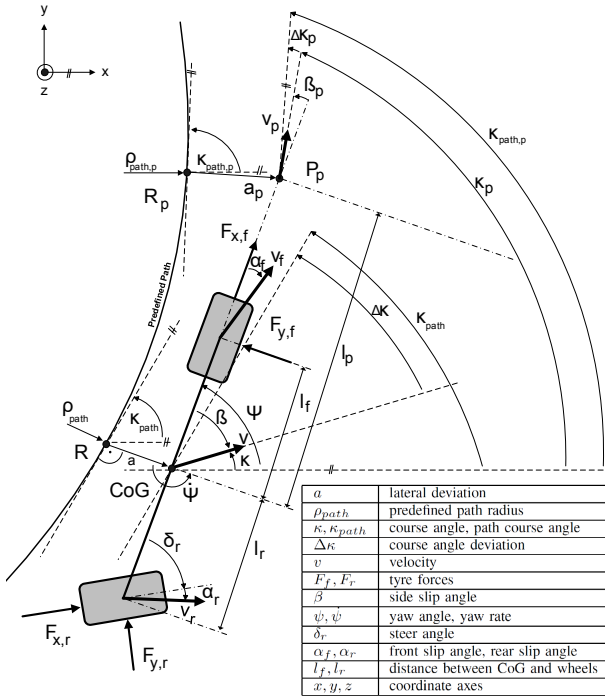


Fig. 3 Single track model with rear axle steering

Based on the presented structure of the mathematical plant model, the modeling of the single parts can be explained. The steering actuator is implemented as a first order delay element with the delay time T_s . The second part is the so-called single track model [7], [19]. This well-known model from the literature is valid for vehicles with front-axle steering. An extension in order to describe the lateral dynamic behaviour of industrial trucks with rear-axle steering has been derived in [1]. It is obtained under the assumption that the CoG of the vehicle is at road level, which neglects the influence of wheel load distributions. The wheels of each axle can thus be combined to one resulting wheel (figure 3). The mathematical description of the lateral dynamic vehicle behaviour is based on the following assumptions:

1. Neglect of longitudinal dynamic forces like traction forces, braking forces and aerodynamic drag forces
2. Constant or only slowly changing vehicle speed
3. Small steering angles, slip angles and side slip angles

The first two assumptions can be made since the lateral vehicle dynamic motion basically changes faster than the

longitudinal dynamic motion. The third assumption limits the range of validity of the model. It describes the lateral vehicle dynamic behaviour only with restricted accuracy at the limits of driving physics. In order to be able to use the plant model for the design of the lateral controller, the model equations have to be extended to describe the relative motion of the vehicle with respect to the path. Finally, the plant model can be given in state space representation (equation 1), where $x = [\beta, \dot{\psi}, \Delta\kappa, a_p, \delta_r]^T$ describes the state vector of the system and $u = [\delta_{set}, \chi_p]^T$ represents the vector of its input signals. These are the steering angle setpoint, calculated by the lateral controller (control signal), as well as the curvature of the predefined path, considered as disturbance variable. Using this information and the state space model (equation 1) the transfer functions of the plant can be specified. They are given in equations (2) and (3) and are used for the classical model-based control design of the 2DoF controller in section 3. Table 1 provides an overview of the associated values of the model parameters used in equations (1) - (3) for two vehicle variants. The Linde E30 is used in the first training step of the RL controller and represents the nominal vehicle variant. The Linde E80 is a larger industrial truck variant and is used to investigate the adaptability of the presented approaches (second training step). A detailed derivation of the plant model, a validation using real measurement data as well as an extensive analysis with respect to fluctuating parameters during operation is given in [1].

$$\begin{bmatrix} \dot{\beta} \\ \ddot{\psi} \\ \Delta\dot{\kappa} \\ \dot{a}_p \\ \dot{\delta}_r \end{bmatrix} = \begin{bmatrix} -\frac{c_f+c_r}{m \cdot v} & -\frac{c_r \cdot l_r + c_f \cdot l_f}{m \cdot v^2} + 1 & 0 & 0 & \frac{c_r}{m \cdot v} \\ -c_r \cdot l_r + c_f \cdot l_f & -\frac{c_r \cdot l_r^2 + c_f \cdot l_f^2}{m \cdot v^2} & 0 & 0 & \frac{c_r \cdot l_r}{J_z} \\ J_z & J_z \cdot v & 0 & 0 & \frac{J_z}{m \cdot v} \\ -\frac{c_f+c_r}{m \cdot v} & \frac{c_f \cdot l_f - c_r \cdot l_r}{m \cdot v^2} & 0 & 0 & 0 \\ 0 & -l_p & v & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T_s} \end{bmatrix} \cdot \begin{bmatrix} \beta \\ \dot{\psi} \\ \Delta\kappa \\ a_p \\ \delta_r \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & v \\ 0 & 0 \\ \frac{1}{T_s} & 0 \end{bmatrix} \cdot \begin{bmatrix} \delta_{set} \\ \chi_p \end{bmatrix} \quad (1)$$

Table 1 Vehicle parameters of the Linde E30 and E80 [26]

Parameters	Linde E30	Linde E80
m	4981 kg	15720 kg
l	1.665 m	2.400 m
c_f	62000 N/rad	62000 N/rad
c_r	122000 N/rad	122000 N/rad
l_f	0.858 m	1.181 m
l_r	0.807 m	1.219 m
l_p	1.5 m	1.5 m
J_z	3624 kgm ²	26490 kgm ²
T_s	0.2 sec	0.2 sec

$$G_{a,\chi}(s) = \frac{v^2}{s^2} \quad (2)$$

$$G_{a,\delta}(s) = K \cdot \frac{s^2 + b_1 \cdot s + b_0}{s^2 \cdot (s + \frac{1}{T_s}) \cdot (s^2 + a_1 \cdot s + a_0)} \quad (3)$$

with

$$K = -\left(\frac{-c_r \cdot J_z + l_p \cdot c_r \cdot l_r \cdot m}{J_z \cdot m \cdot T_s}\right), b_0 = \frac{c_r \cdot c_f \cdot (l_r + l_f)}{-c_r \cdot J_z + l_p \cdot c_r \cdot l_r \cdot m},$$

$$b_1 = \frac{c_r \cdot c_f \cdot (l_r \cdot l_f + l_f^2) - l_p \cdot c_r \cdot c_f \cdot (l_f + l_r)}{-v \cdot (-c_r \cdot J_z + l_p \cdot c_r \cdot l_r \cdot m)},$$

$$a_0 = \frac{c_r \cdot l_r - c_f \cdot l_f}{J_z} + \frac{c_r \cdot c_f \cdot (l_f + l_r)^2}{J_z \cdot m \cdot v^2},$$

$$a_1 = \frac{c_f + c_r}{m \cdot v} + \frac{c_r \cdot l_r^2 + c_f \cdot l_f^2}{J_z \cdot v}$$

2.2. 2DoF control structure

Figure 4 presents the structure of the 2DoF control concept which can be used to compensate the influence of the disturbance variable, i.e. the path curvature [7] - [11]. The control signal δ_{set} is formed by superposition of two signal components. The first part (δ_{FFC}) is calculated by a FFC that uses the detailed path information, which are available in advance [18]. It determines the control signal in dependence of the path curvature χ_p in the current reference point R_p based on a simplified and linearized model of the plant (subsection 2.1) [7]. The FBC calculates the second component (δ_{FBC}) of the control signal. Its task is to stabilize the plant and to compensate the lateral deviation a_p caused by model inaccuracies and other disturbances. In addition to the described advantages of this control concept, it has a decisive disadvantage with regard to the task of automation of a heterogeneous fleet. The FFC is not adaptive to different vehicle variants. Although the FBC can compensate for minor variances during operation, an adaptation to another truck variant is not possible with this control approach.

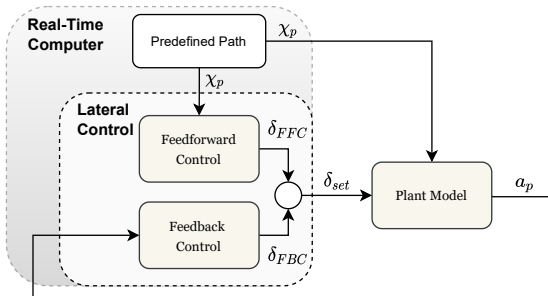


Fig. 4 Structure of the 2DoF control concept

2.3. Proposed AI-based control structure

Figure 5 depicts the control structure of the proposed AI-based control concept. It is based on the RLCCPK concept given in [1]. In order to take into account the influence of the disturbance variable, the structure of the

RLCCPK control concept is extended in analogy to the 2DoF concept. Since the path is defined in advance and stored on the real-time computer (section 1), the path curvature in the reference point R_p can be used as a priori knowledge. Thus, this information (χ_p) is provided to the lateral controller as an additional input signal.

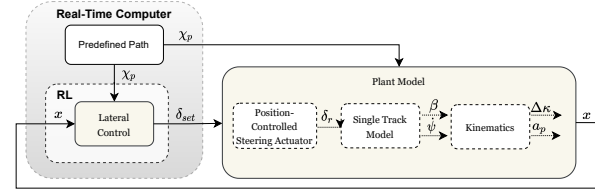


Fig. 5 Structure of the vehicle guidance system

The calculation of the control signal δ_{set} is based on the current system state x on the one hand as well as on the current path curvature in the reference point R_p on the other hand. With this new control structure, the advantages of the RLCCPK and the 2DoF control concept can be combined. It results in a new approach that is able to adapt to different vehicle variants taking into account the a priori plant knowledge and to compensate the influence of the varying path curvature during operation.

3. DESIGN OF THE 2DOF CONTROLLER

In section 2 it was pointed out that the curvature of the path χ_p in the current reference point R_p can be regarded as a disturbance variable of the lateral vehicle guidance system. Since the path is predefined and stored on the real-time computer, this a priori knowledge offers the possibility to reduce the influence of the varying path curvature during operation by means of a disturbance rejection [11]. Assuming that the mathematical model describes the controlled system accurately, the influence of the disturbance variable can completely be compensated with a suitable definition of the FFC ($G_{FFC}(s)$).

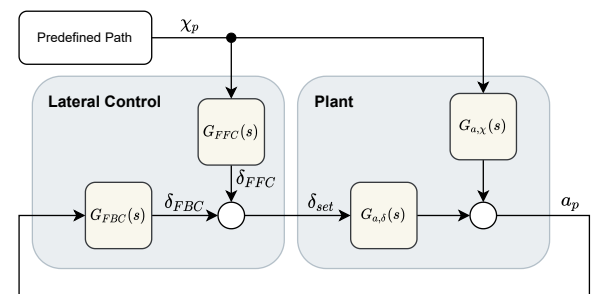


Fig. 6 Structure of the 2DOF-control concept

Figure 6 shows the structure of the 2DoF control concept. Its design is based on the disturbance transfer function ($G_{a,\chi}(s)$) and the control transfer function ($G_{a,\delta}(s)$) of the plant, given in the equations (2) and (3) in section 2. In this case, $G_{a,\chi}(s)$ describes the effects of the path curvature χ_p on the system's output a_p . $G_{a,\delta}(s)$ characterizes the dynamic behaviour of the vehicle and the steering actuator. Based on this, the following calculation of $G_{FFC}(s)$ can be derived:

$$G_{a,\chi}(s) - G_{FFC}(s) \cdot G_{a,\delta}(s) = 0$$

$$\Rightarrow G_{FFC}(s) = \frac{G_{a,\chi}(s)}{G_{a,\delta}(s)} \quad (4)$$

Since the resulting transfer function $G_{FFC}(s)$ (equation 4) has a higher number of zeros than poles, a first order low-pass filter with a small time constant T_{FFC} has to be added. As the FFC does not ensure a precise track guidance by itself, a FBC is used to compensate the occurring lateral deviation a_p . This procedure increases the robustness of the control system with respect to imprecisely known model parameters and stabilize the controlled system.

The FBC is designed using the root locus method in order to achieve a damping of the dominant poles of about $D = 0.7$. A detailed description of the control design using root locus method has already been given in [7], [9], [10]. The associated transfer function ($G_{FBC}(s)$) represents the FBC as a PDT_1 controller (equation 5), where K_{FBC} is the gain factor, T_D the derivative time and T_{FBC} the time constant of a first order low-pass filter. The associated control parameters of the 2DoF controller are given in table 2.

$$G_{FBC}(s) = K_{FBC} \cdot \frac{T_D \cdot s + 1}{T_{FBC} \cdot s + 1} \quad (5)$$

Table 2 Parameters of the 2DoF controller

Control Parameters	Values
T_{FFC}	0.01 sec
K_{FBC}	3.2634
T_D	0.5 sec
T_{FBC}	0.02 sec

4. AI-BASED CONTROL APPROACHES

This section is dedicated to the AI-based control approaches for the automatic track guidance of industrial trucks. At the beginning, the used methodology and the basics of RL are presented. Subsection 4.2 introduces the RLCCPK approach given in [1], since it is used as a comparison control concept in section 5. Finally, the proposed RLDC approach is discussed in subsection 4.3 that specifically considers the varying curvature of the predefined path during operation.

4.1. Reinforcement Learning basics

RL is a well-known approach in the domain of control systems [20], [21], [22]. It is assigned to the methods of direct neural control, since AI acts as a controller and calculates the control signal by itself. The training of the RL controller takes place in closed-loop operation and is done in analogy to the human learning process.

Experiences are built up by interacting with the system. The principle of the RL process is displayed in figure 7 and essentially consists of three blocks. The undermost block (vehicle) represents the controlled system, in this case the industrial truck. Its current state Φ_k is provided to the RL controller. This block describes the lateral controller that calculates the control signal u_k in order to affect the controlled system. The third block (reward function) evaluates the control signal u_k based on the current state Φ_k and the following state Φ_{k+1} , in form of a feedback, called reward r_k . It is a measure of control quality. In analogy to the human learning process, the control strategy is adapted in order to optimize the reward.

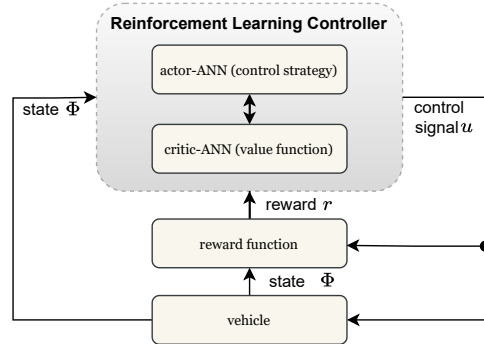


Fig. 7 Principle of Reinforcement Learning

The described basic idea of RL can be implemented using different methods. In this paper the TD3 algorithm is used, which is an extension of the Deep Deterministic Policy Gradient (DDPG) algorithm [21]. The TD3 algorithm is well suited for the application of automatic track guidance based on two main reasons. On the one hand, the RL controller is able to calculate a value continuous control signal which is important for a smooth vehicle track guidance. On the other hand, the training process proves to be more stable compared to the DDPG algorithm due to additional target-nets [17]. TD3 is a so-called Actor-Critic method that uses separate memory structures to differ between the control strategy $\mu(\Phi)$ (actor-ANN) and the value function $Q(\Phi, u)$ (critic-ANN). $Q(\Phi, u)$ is a function to calculate the expected cumulative reward \hat{r} , based on its input signals Φ and u and represents the knowledge of the plant. This means, it evaluates the expected reaction of the controlled system with respect to the calculated control signal in the current system state. The optimization of the parameters ϕ of the critic-ANN is done by supervised learning, based on the obtained reward [23], [24]. The task of the actor-ANN consists of calculating the control signal u_k in dependence of the current system state Φ_k and is indicated as a function of the actor-ANN parameters θ . The optimization of this parameters (θ) should be done in order to maximize the output of the critic-ANN and thus the reward. To implement this, a criterion J is defined that describes the start distribution of $Q(\Phi, u)$ [21]. The basic idea is to adjust the parameters of the actor-ANN θ in the direction of the gradient $\nabla_{\theta} J$ [17], [21], [25]. This is done by applying the chain rule with respect to the actor-ANN parameters θ (equation 6):

$$\nabla_{\theta} J \approx \frac{1}{N} \sum_i^N \nabla_u Q(\Phi, u | \phi) |_{\Phi=\Phi_i, u=\mu(\Phi_i)} \nabla_{\theta} \mu(\Phi | \theta) |_{\Phi=\Phi_i} \quad (6)$$

The observation vector Φ , reflecting the state of the system, is depending on the chosen methodology. Whether the disturbance variable is taken into account (RLCDC) or not (RLCCPK), the observation vector is composed differently (subsections 4.2 and 4.3).

4.2. RLCCPK approach

The RLCCPK approach given in [1] does not consider the influence of the disturbance variable. The used observation vector Φ is formed similar to the state vector x of the model described in section 2 and is given in equation (7):

$$\Phi = x = [\beta, \dot{\psi}, \Delta\kappa, a_p, \delta_r]^T \quad (7)$$

The behaviour of the RL controller can be specified by the definition of the reward function. [1] demonstrates that closed-loop behaviour of optimal state control can be approximated by choosing the reward function r_k in analogy to the quadratic cost function of classical LQR [27]. In this application the used reward function of the RLCCPK is defined to focus on minimizing the lateral deviation a_p of the vehicle with respect to the path. Therefore, the weighting factor of $a_{p,k}^2$ is chosen significantly larger than the weightings of the other signals (equation 8).

$$r_k = -(\beta_k^2 + \dot{\psi}_k^2 + \Delta\kappa_k^2 + 10000 \cdot a_{p,k}^2 + \delta_{r,k}^2 + 5 \cdot \delta_{set,k}^2) \quad (8)$$

4.3. Proposed RLCDC approach

In order to compensate the influence of the varying path curvature in the reference point R_p during operation, the observation vector Φ (equation 7) of the RLCCPK is extended by the disturbance variable χ_p , leading to Φ_{ext} of the RLCDC approach (equation 9):

$$\Phi_{ext} = \begin{bmatrix} \Phi \\ \chi_p \end{bmatrix} = [\beta, \dot{\psi}, \Delta\kappa, a_p, \delta_r, \chi_p]^T \quad (9)$$

Since the signals of the observation vector form the inputs of the actor-ANN and the critic-ANN of the RL controller, the structure of these networks has to be adjusted. A further neuron is integrated in the input layers of the ANN, in order to process the information of the enlarged observation vector. Figure 8 depicts a simplified representation of the structure of the actor-ANN (left) and the critic-ANN (right). In the first hidden layer of both fully connected feed-forward ANN, 400 neurons are inserted. Therefore, the extension of the input layer with an additional neuron results in a large number of further ANN parameters.

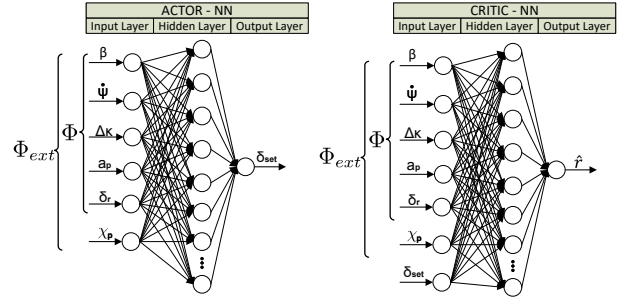


Fig. 8 Simplified representation of the extended ANN structure of the RLCDC approach

In order to compare the different RL control concepts (RLCCPK and RLCDC) with each other, the reward function given in equation (8) is used for the RLCDC approach as well.

5. CONTROL DESIGN AND SIMULATION RESULTS

In this section, the simulation results of both RL approaches (RLCCPK and RLCDC) and the 2DoF control concept are presented and compared with each other. Subsection 5.1 focuses on the results after the first training step of the RL approaches (pre-training). This first training step is performed using the model parameters of a nominal industrial truck variant (Linde E30). The 2DoF controller is designed for this industrial truck variant as well (table 1). Subsequently, the adaptability of the RL concepts to another vehicle variant, such as the Linde E80 will be discussed in subsection 5.2. For this purpose, the second training step (fine tuning) is performed based on the pre-trained controllers (subsection 5.1). Since the 2DoF control concept is not adaptive, the simulation results of the Linde E80 are also presented using the 2DoF controller designed for the Linde E30. Both training steps are carried out in simulation using the model described in section 2, considering the scenario given in figure 9.

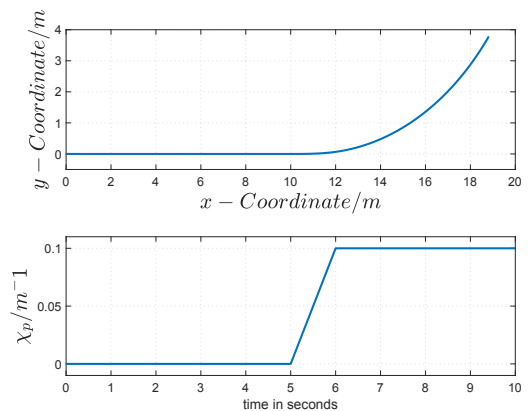


Fig. 9 Test scenario

The upper part of the figure displays the course of the predefined path [0-20m]. The path initially runs as a straight line [0-10m] and merges into a curve with a constant curve radius ρ_{path} . The transition between the mentioned segments is realized as a clothoid [10-12m], where the radius is linearly reduced until it reaches the

final curve radius [12-20m]. Since the control concepts refer to a constant velocity of $v = 2\text{m/s}$ during the entire test scenario, the path curvature can be calculated. It is shown in the lower part of figure 9 and is applied to the system as a disturbance variable χ_p (section 2). The industrial truck starts with an initial lateral deviation of the preview point of $a_p = 0.2\text{m}$, i.e. offset from the path.

5.1. Simulation results after the first training step

This subsection compares the RLCCPK, RLDC after completion of the first training step and the 2DoF controller. The simulation results of the nominal vehicle variant (Linde E30) are presented. It shall be shown, how the different control concepts can deal with the scenario given in figure 9 and compensate the influence of the disturbance variable. Since RLCCPK does not consider the varying path curvature during operation, this approach is trained without disturbance signals in all training epochs. In order to take into account occurring path curvatures during operation, the structure of the ANN of the RLDC is adjusted as discussed in subsection 4.3. The training process of the RLDC controller is divided into several epochs, each of them with a different disturbance value within the range of $[-0.3 \leq \chi_p \leq 0.3]$.

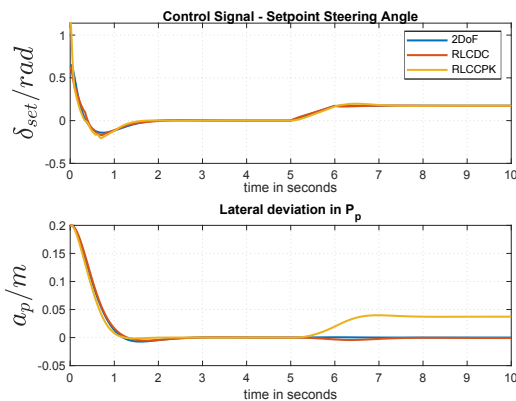


Fig. 10 Steering angle and lateral deviation (Linde E30)

Figure 10 shows the simulation results of all presented control concepts. In the upper part of the figure, the time courses of the control variable (δ_{set}) is depicted. The controlled variable (a_p) is illustrated below. Obviously, all three concepts are comparable in the range [0sec - 5sec]. The lateral deviation of the RLCCPK differs from the other control concepts in the rear part [5sec - 10sec]. This is due to the fact that the path curvature is applied to the system and not taken into account by the RLCCPK. Obviously, the extension of the RL approach (RLDC) almost completely compensates the influence of the disturbance variable and leads to a steady state accuracy. The extension of the input layer in the ANN of the RLDC in combination with the high number of neurons of the first hidden layer, leads to a more complex ANN with a large number of additional ANN parameters. This results in a higher degree of freedom with respect to the design and improves the control quality by compensating the influence of the disturbance variable. However, it has a negative effect on the training efficiency, since sig-

nificantly more optimization steps have to be performed. This can be seen by comparing the optimization steps in the first training step of RLCCPK and RLDC (first and third row in table 3).

5.2. Simulation results after the second training step

In this subsection, the adaption of the control concepts to another industrial truck variant is investigated. To avoid starting the entire training process for another vehicle variant from scratch the pre-trained RL controllers of subsection 5.1 are used as starting point for the second training step. Both controllers have to be adapted within the second training step (fine tuning) to the actual industrial truck variant, in this case the Linde E80 and the associated vehicle parameters (table 1). By this method, the number of optimization steps can be significantly reduced compared to a training that has to be started from scratch. This can be illustrated by comparing the required optimization steps in the first and second training step within a control approach (first and second row or third and fourth row of table 3).

Table 3 Training efficiency

Concept	Training	Optimization steps
RLCCPK	1st step E30	116822
RLCCPK	2nd step E80	14000
RLDC	1st step E30	295693
RLDC	2nd step E80	26000

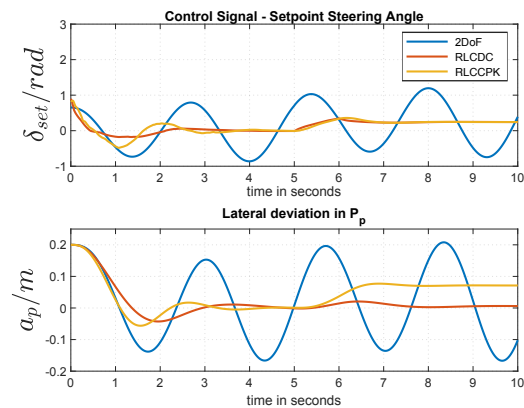


Fig. 11 Steering angle and lateral deviation (Linde E80)

Figure 11 shows the performance of the three control concepts for the presented scenario. It can be seen that the 2DoF controller designed for the Linde E30 is not able to stabilize the Linde E80. This is due to the fact that the dynamics of the two vehicle variants are significantly different, which affects the design of both the FFC and the FBC. The RL controllers are adapted to the changed vehicle variant in the second training step and are capable of stabilizing the vehicle. The RLDC is still able to compensate the influence of the disturbance variable, which can be seen in the range [5sec - 10sec]. Thus, it shows clear advantages over the RLCCPK approach which can't provide steady state accuracy. Moreover, the closed control loop behaviour using the RLDC controller results in a significantly higher damping.

6. CONCLUSION

This paper presents a new AI-based control approach for the automatic track guidance of industrial trucks. By separating the training process into two steps, existing a priori knowledge regarding the controlled system can be integrated during the training. In the first training process, the RL controller's experience is built up in simulation using a linear model and the parameters of a nominal, average vehicle variant. Since the basic dynamics are comparable for all truck variants, the experience buildup does not have to be performed over and over again from scratch. Therefore, based on the pre-trained controller, an adaptation to other vehicle variants can be performed by fine-tuning the controller parameters in a second training step. By extending the observation vector and the ANN used in the RL controller, a compensation of the influence of the path curvature is possible. Thus, the control quality of the concept can be improved and a stable control loop behaviour for different industrial truck variants can be ensured in the investigated scenarios. With the new control concept RLDC, the advantages of the other presented control concepts can be combined. The adaptability with regard to new industrial truck variants of the RLDC is combined with the possibility of compensating the influence of disturbance variables of the 2DoF control concept. Finally, it should be mentioned that the RL concepts have a significant disadvantage compared to the 2DoF approach. In this configuration of the RL control approaches, all state variables of the system have to be available to the controller, whereas the 2DoF concept only requires the output variable of the plant.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the Bavarian Ministry of Economic Affairs, Regional Development and Energy for the financial support of the project *Cooperative Autonomous Intralogistic Systems* under Grant No. DIK-1910-0016 and the industrial project partner: Linde Material Handling GmbH.

REFERENCES

- [1] Sauer et al.: Automatic track guidance of industrial trucks using self-learning controllers considering a priori plant knowledge, 5th ICCAD, 2021.
- [2] Tan et al.: Development of an Automated Steering Vehicle Based on Roadway Magnets - A case Study of Mechatronic System Design, IEEE/ASME TOM, Vol. 4, 1999.
- [3] Li et al.: Path-Following Control for Self-driving Forklifts based on Cascade Disturbance Rejection with Coordinates Reconstruction, 39th CCC, 2020.
- [4] Tamba et al.: Trajectory Generation of an Unmanned Forklift for Autonomous Operation in Material Handling System, SICE AC, 2008.
- [5] Ritzer et al.: Advanced Path Following Control of an Overactuated Robotic Vehicle, IEEE Intelligent Vehicle Symposium (IV), 2015.
- [6] Mohammadi et al.: Model Predictive Motion Control of Autonomous Forklift Vehicles with Dynamics Balance Constraint, 14th ICCARV, 2016.
- [7] Zindler et al.: Querdynamische Fahrzeugführung zur reproduzierbaren Erprobung von Sicherheitssystemen. *at-Automatisierungstechnik* 60(2), Oldenbourg-Verlag, S. 61-73, 2012.
- [8] Hahn et al.: Two-Degrees-of-Freedom Lateral Vehicle Control using Nonlinear Model Based Disturbance Compensation, 8th IFAC-AAC, 2016.
- [9] Heinlein et al.: Control methods for automated testing of preventive pedestrian protection systems, *International Journal of Vehicle Systems Modelling and Testing*, Vol. 10, No. 2, 2015.
- [10] S. Hahn: Methoden zur nichtlinearen modellbasierten Spurführung benutzerdefinierter Punkte an der Fahrzeugfront, PhD-thesis, 2017.
- [11] I.M. Horowitz: *Synthesis of Feedback Systems*. Academic Press, New York, 1963.
- [12] Landau et al.: *Adaptive Control: Algorithms, Analysis and Applications*, Springer, London, 2013.
- [13] K.J.Aström, B. Wittenmark: *Adaptive control*, Mass: Addison-Wesley, Reading, 2. Edition, 1995.
- [14] W. Levine and T. Sawa: *The Control Handbook*, CRC PRESS, IEEE Press, 1996.
- [15] Sallab et al.: End-to-End Deep Reinforcement Learning for Lane Keeping Assist, 30th Conference on NIPS, 2016.
- [16] Havenstrom et al. "Proportional integral derivative controller assisted reinforcement learning for path following by autonomous underwater vehicles." arXiv preprint arXiv:2002.01022, 2020
- [17] Fujimoto et al.: Addressing Function Approximation Error in Actor-Critic Methods, ICML, 2018.
- [18] I. Söhnitz: *Querregelung eines autonomen Strassenfahrzeugs*. Fortschr.-Ber. VDI Reihe 8, Nr. 882. VDI Verlag Düsseldorf, 2001.
- [19] H. Pacejka: *Tyre and vehicle dynamics* (2nd ed.), Oxford: Butterworth-Heinemann, 2006.
- [20] M. Vogt: An overview of deep learning techniques, *at - Automatisierungstechnik* 66 (9), pages 690-703, Oldenbourg Wissenschaftsverlag, 2018.
- [21] Lillicrap et al.: Continuous Control with Deep Reinforcement Learning, ICLR, 2016.
- [22] R. Sutton and A. Barto: *Reinforcement Learning: an introduction*. The MIT Press, 2018.
- [23] M. Hagan et al.: *Neural Network Design*, Martin Hagan, 2014.
- [24] K. Gurney: *An introduction to neural networks*, UCL Press Limited, 1997.
- [25] D. Silver et al.: Deterministic Policy Gradient Algorithms. ICML, 2014.
- [26] LINDE MATERIAL HANDLING. Homepage Linde Material Handling. Accessed on 22.02.2022.
- [27] Ichikawa et al.: Neural Network Application for Direct Feedback Controllers, *IEEE Transactions on neural networks*, VOL. 3, NO. 2, 1992.