

# Direct Homography Control for Vision-Based Platooning

Alexander Schaub<sup>1</sup>, Ricardo de Castro<sup>1</sup> and Darius Burschka<sup>2</sup>

**Abstract**—This paper introduces a vision-based controller for automatic vehicle following, also known as 2-vehicle platooning. A direct homography controller is applied to calculate the motion demand for an autonomous vehicle from only the data of a monocular camera. The direct control without an intermediate step to a Cartesian representation increases the robustness of the scheme. A robustness analysis of the closed loop controller is provided using the parameter space approach. Furthermore, the direct homography controller is extended by an estimation of the absolute angular difference to the goal position, which then enables the estimation of the position error. The proposed homography-based position estimation is tested on rendered camera images for better evaluation of the underlying error and the platooning controller is verified in simulation. Finally, the results of both are presented.

## I. INTRODUCTION

For decades, cameras have been a widely used sensor for autonomous mobile robots or autonomous vehicles [1]. Nowadays, a huge variety of driver assistance systems uses cameras for lane or traffic sign detection, automatic headlights or even automatic braking systems, which utilize approaches based on images processing like the so-called '6D-Vision' [2]. The choice for cameras is made due to various advantages like passivity, energy consumption, information density and cost efficiency. Object or light detection systems usually only require a single camera, but planning and executing autonomous motions based on data from a monocular camera is challenging, since the perception is 2D and sizes are only perceived up-to-scale.

Even classic image-based visual servoing approaches, where the measurements and the control error are directly expressed in image plane coordinates, have to measure or estimate the depth of feature points [3],[4]. In order to build an image-based controller, which provides outputs that are not only up-to-scale, a possible solution is the utilization of the homography principle. For this the target view must be known and the feature points must lie on a planar surface [5]. Usually, the homography matrix is decomposed to receive the relative translation and rotation from the current camera pose to the target pose in Cartesian coordinates.

This decomposition is utilized in [6] for the position estimation of a preceding vehicle in a platooning scenario, whereas the controller works in a Cartesian reference frame spanned by the preceding vehicle. In [7] homography decomposition

is utilized for estimating the leading vehicle's position together with its velocity.

One drawback of the decomposition is that it provides two physically feasible solutions from which one has to be chosen. Another shortcoming is that the transformation to Cartesian space impairs a direct use of the image data, which would increase the robustness, as a Cartesian representation requires additional processing with error prone parameters, and decrease the computational costs. An attempt to avoid the decomposition is made in [8], where the motion of a unicycle robot is planed and controlled by three motion primitives. For those motion primitives a controller is designed utilizing the entries of the homography matrix, but for certain motions the decomposition still has to be carried out. A similar approach using two steps is described in [9]. A decomposition-free visual servoing control for a 6-DOF robot is described [10] with restrictions regarding the relative position of the goal pose to the plane of the image features. The same principle is utilized in [11] for positioning an underwater vehicle.

This work utilizes the approach of [12], where a direct homography control (DHC) is introduced for highly maneuverable vehicles like e.g. the German Aerospace Center's ROboMObil [13]. Here, only the entries of the homography matrix are used in the feedback loop without any decomposition or restrictions regarding the target pose.

This article has three main contributions. First the direct homography based controller is improved by an a priori estimation of the angular difference to the goal position. Second it is shown that the improved DHC can be transformed to a position estimator for planar motions using only entries from the homography matrix. Finally, the DHC is integrated in a platooning scheme for highly maneuverable vehicles with a monocular camera being the only perception sensor.

The remainder of this article is organized as follows. Section 2 first reviews the homography principle leading to the DHC, then introduces the a priori determination of the angular state and finally describes the step to the position estimation. In Section 3 the DHC-based platooning controller is described and a robustness analysis is carried out. Section 4 provides the results for the position estimation and the platooning, while Section 5 concludes the paper and gives a brief outlook.

## II. HOMOGRAPHY MODELING

The basic homography principle is depicted in Figure 1. There, the target pose of a camera's optical center denoted as  $O^*$  with  $\chi^* = [X_C^*, Y_C^*, Z_C^*]$  being the camera centered Cartesian coordinate system. A planar surface  $\Pi$  containing single points  $\mathbf{p}_i$  is observed by the camera with  $\mathbf{n}$  being the

<sup>1</sup>Institute of System Dynamics and Control, German Aerospace Center, 82234 Wessling, Germany `firstname.lastname@dlr.de`

<sup>2</sup>Institut für Informatik VI Technische Universität München 85748 Garching bei München, Germany `burschka@cs.tum.edu`

unit normal vector to  $\Pi$  and  $d$  being the distance between the image plane  $I^*$  and  $\Pi$ . Moreover,  $\mathbf{m}^*$  denotes the intersection of a ray from the optical center to a point  $\mathbf{p}_m$  with  $I^*$ .

If the camera is moved via a translation  $\mathbf{t} \in \mathbb{R}^3$  and Rotation  $\mathbf{R} \in \mathbb{R}^{3 \times 3}$  to a second view at point  $O$ , the relation between  $\mathbf{m}$ , which is the projection of  $\mathbf{p}_m$  into image plane  $I$ , and  $\mathbf{m}^*$  could then be written as following:

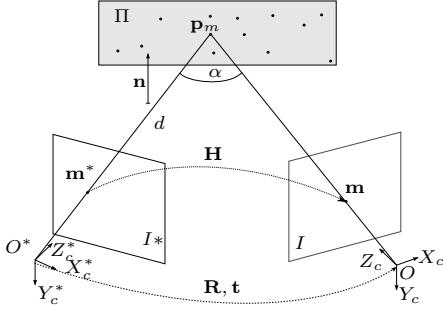


Fig. 1. Homography Principle

$$\mathbf{m} = \mathbf{R}\mathbf{m}^* + \mathbf{t} \frac{1}{d} \mathbf{n}^T \mathbf{m}^* \Leftrightarrow \mathbf{m} = (\mathbf{R} + \frac{1}{d} \mathbf{t} \mathbf{n}^T) \mathbf{m}^* \quad (1)$$

The matrix denoting the relation between both projected points is called the homography matrix  $\mathbf{H}$ . This relation is up to scale  $\eta^i m^i = \hat{m}^i, i \in \{*, \}$ , as the translation  $\mathbf{t}$  is scaled by the distance  $d$ . Nevertheless, this factor can be determined, if at least four feature points are available to compose  $\mathbf{H}$  [5]:

$$\mathbf{m} = \mathbf{H}\mathbf{m}^* \quad (2)$$

#### A. Direct Homography Control

In order to calculate the rotation  $\mathbf{R}$  and translation  $\mathbf{t}$  a decomposition of the homography matrix could be carried out that provides 4 solutions, whereas only 2 are physically feasible. The correct solution can be determined then via the estimated normal vector  $\mathbf{n}$  [14]. Nevertheless, the transformation into the Cartesian space should be omitted in order to avoid the ambiguity, to increase the robustness and to decrease the computational costs.

The basic idea of a direct homography control is to derive the controller output directly from the homography matrix, which is an implicit representation of the error to the 3D target pose. The advantage of  $\mathbf{H}$  is that it can be derived only from 2D feature point measurements and therefore requires only a monocular camera. First, since the goal of this work is to control an autonomous vehicle or mobile robot, the state space could be reduced to two translational and one rotational degrees of freedom, namely the camera's  $X_c$ - and  $Z_c$ -axis and the rotation around the height axis  $Y_c$ . Hence,  $\xi$  is a 3 dimensional vector and the Homography matrix can be simplified to:

$$\mathbf{H}(\alpha, \mathbf{t}) = \begin{pmatrix} \cos(\alpha) + \frac{t_1 n_1}{d} & \frac{t_1 n_2}{d} & \sin(\alpha) + \frac{t_1 n_3}{d} \\ \frac{t_2 n_1}{d} & \frac{t_2 n_2}{d} + 1 & \frac{t_2 n_3}{d} \\ -\sin(\alpha) + \frac{t_3 n_1}{d} & \frac{t_3 n_2}{d} & \cos(\alpha) + \frac{t_3 n_3}{d} \end{pmatrix} \quad (3)$$

Here,  $n_1$  is the first entry of  $\mathbf{n} = [n_{X_c}, n_{Y_c}, n_{Z_c}]^T$  and denotes the value in  $X_c$  direction of the camera, while the same indexes hold for elements of  $\mathbf{t}$ , so that the state  $\xi$  can be expressed as:

$$\xi = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{t} \\ \alpha \end{pmatrix} = \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta \phi \end{pmatrix} \quad (4)$$

The vector  $\xi$  denotes the relative position to the goal pose in the vehicle's coordinate system. The notation is chosen in that way, as the camera's coordinate system does not coincide with the coordinates of the vehicle. The utilization of  $\mathbf{H}$  in a control scheme is eased by stacking the single entries in a vector  $\mathbf{h} \in \mathbb{R}^9$ :

$$\mathbf{H} = (\mathbf{H}_1^T \quad \mathbf{H}_2^T \quad \mathbf{H}_3^T)^T \Rightarrow \mathbf{h} = (\mathbf{H}_1 \quad \mathbf{H}_2 \quad \mathbf{H}_3)^T \quad (5)$$

In order to see how position changes affect the entries of the homography matrix a Jacobian matrix  $\mathbf{J}$  is obtained by combining (5) and the time derivative of (4):

$$\dot{\mathbf{h}} = \begin{pmatrix} \frac{\delta \mathbf{H}_1}{\delta \xi} & \frac{\delta \mathbf{H}_2}{\delta \xi} & \frac{\delta \mathbf{H}_3}{\delta \xi} \end{pmatrix}^T \frac{\delta \xi}{\delta t} \quad (6)$$

The entries of  $\dot{\xi}$  are the relative lateral velocity  $\delta v_Y$ , the relative longitudinal velocity  $\delta v_X$  and the relative yaw rate  $\delta \dot{\psi}$ . In order to prepare for the case of a moving goal position, a distinction is made between the motion of the controlled vehicle  $\mathbf{u} = [v_X, v_Y, \dot{\psi}]^T$  and the velocities of the target pose or the preceding vehicle  $PV$  in the platooning case that is depicted in Figure 2.

$$\begin{aligned} \dot{t}_1 &= -\delta v_Y = -(v_Y - v_{Y,PV}) \\ \dot{t}_3 &= \delta v_X = v_X - v_{X,PV} \\ \dot{\alpha} &= \delta \dot{\psi} = \dot{\psi} - \dot{\psi}_{PV} \end{aligned} \quad (7)$$

For a standard positioning task the velocity vector of the target pose  $\mathbf{u}_{PV} = [v_{(X,PV)}, v_{(Y,PV)}, \dot{\psi}_{PV}]^T$  is zero. For the platooning approach the non-changing parameters  $\mathbf{n}$  and  $d$  should be known or roughly estimated. Now, the Jacobi matrix can be calculated as the time derivative as:

$$\dot{\mathbf{h}} = \underbrace{\begin{pmatrix} 0 & -\frac{n_1}{d} & -\sin(\alpha) \\ 0 & -\frac{n_2}{d} & 0 \\ 0 & -\frac{n_3}{d} & \cos(\alpha) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{n_1}{d} & 0 & -\cos(\alpha) \\ \frac{n_2}{d} & 0 & 0 \\ \frac{n_3}{d} & 0 & -\sin(\alpha) \end{pmatrix}}_{\mathbf{J}(\alpha)} [\delta v_x, \delta v_y, \delta \dot{\psi}]^T \quad (8)$$

#### B. Determining $\alpha$

The angular error  $\alpha$  is not directly measurable from the homography matrix. In [12] a linearization around the goal position was utilized with good results, since the error converges to zero and the small-angle approximation holds. Nevertheless, the real value of  $\alpha$  or an accurate measurement would improve the results.

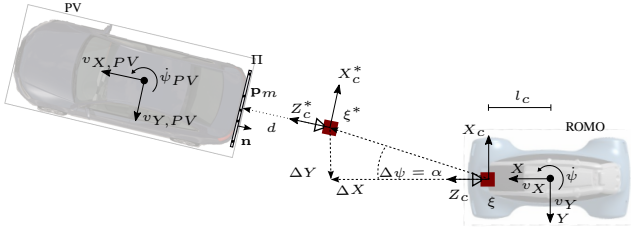


Fig. 2. Platooning using the Homography Principle

First, we notice that some elements of the Homography vector (5), namely  $(h_{12}, h_{21}, h_{22}, h_{23}, h_{32})$ , are always static and do not play a role in our problem setup. This observation enables us to reduce  $\mathbf{h}$  from Equation (5) to a 4-dimension vector  $\underline{\mathbf{h}} \in \mathbb{R}^4$

$$\underline{\mathbf{h}} = (h_{11} \ h_{13} \ h_{31} \ h_{33})^T \quad (9)$$

This reduced homography vector can then be used to generate  $q = \binom{4}{3} = 4$  sets of equations for the three dimensional positioning problem. It should be noted that for a normal vector  $\mathbf{n}$  of rank 1 only 2 of the sets are feasible.

Based on these considerations, the reduced homography dynamics can now be expressed e.g. by a set consisting of  $\underline{\dot{\mathbf{h}}}^1 = (\dot{h}_{11} \ \dot{h}_{13} \ \dot{h}_{31})^T$ :

$$\underline{\dot{\mathbf{h}}}^1 = (\dot{h}_{11} \ \dot{h}_{13} \ \dot{h}_{31})^T = \underline{\mathbf{J}}^1(\alpha)(\mathbf{u} - \mathbf{u}_{PV}) \quad (10)$$

where  $\underline{\mathbf{J}}(\alpha)^1$  is a reduced version of the Jacobian matrix (8) (and obtained by extracting the 1th, 3th and 7th rows of  $\mathbf{J}$ ). Solving the  $q \in \{2, 4\}$  feasible sets, like e.g. Equation (10), analytically for a measured  $\mathbf{h}_m$  will provide  $q$  solutions for the state vector  $\xi^i, i \in [1, \dots, q]$ :

$$\xi^i = f^i(\underline{\mathbf{h}}_m^i, \alpha) \quad (11)$$

For a known  $\alpha$  the solutions of the state vectors are identical:

$$\xi^i = \xi^j, i, j \in [1, \dots, q]$$

This relation could be utilized for a simple minimization procedure to estimate the true alpha, which has to lie in the physically feasible range of  $\alpha \in ]-\pi, \pi[$ :

$$\hat{\alpha} = \hat{g}(\mathbf{h}) : \min_{\alpha} \sum_{i=1}^{q-1} \sum_{j=2}^q \|f^i(\underline{\mathbf{h}}_m^i, \alpha) - f^j(\underline{\mathbf{h}}_m^j, \alpha)\| \quad (12)$$

Since the functions  $f$  are analytically known (compare [12]) and consist only of basic arithmetic operations, solving the optimization is computationally simple. Zero-order optimization methods like line searches are sufficient to get an accurate estimate  $\hat{\alpha}$ .

### C. Position Estimation

Now being able to estimate  $\alpha$  the approach could be utilized to estimate the relative position. Therefore, not the Jacobian matrix that provides the time derivatives, see Equation (8), is needed, but a function or mapping matrix  $\mathbf{F}$  that provides a  $\Delta\xi$  for a measured  $\Delta\mathbf{h}_m$ .

$$\Delta\mathbf{h}_m = \mathbf{h}_m - \mathbf{h}^* = \mathbf{F} \cdot \Delta\xi \quad (13)$$

The goal value for  $\mathbf{H}$  is the identity matrix, since then the current view is identical to the goal view. Hence,  $\mathbf{h}^*$  is the stacked identity matrix. The solution for elements of  $\Delta\mathbf{h}$  without the angle  $\alpha$  is straight forward like e.g.:

$$\Delta h_{12} = \Delta t_1 \frac{n_2}{d} = -\frac{n_2}{d} \Delta y \quad (14)$$

It is more difficult to approximate the non-linear parts containing  $\cos(\alpha)$  or  $\sin(\alpha)$  as a function dependent on  $\Delta\alpha$ .

$$\begin{aligned} \Delta h_{11} &= \cos \alpha + (t_1^* + \Delta t_1) \frac{n_1}{d} - \cos \alpha^* - t_1^* \frac{n_1}{d} \\ &= \cos \alpha - \cos \alpha^* + \Delta t_1 \frac{n_1}{d} \end{aligned} \quad (15)$$

Now, the subtraction term of the angles can be rewritten as:

$$\begin{aligned} \cos \alpha - \cos \alpha^* &= -2 \sin \frac{\alpha + \alpha^*}{2} \sin \frac{\alpha - \alpha^*}{2} \\ &\stackrel{\alpha^* \rightarrow 0}{\rightarrow} -2 \sin \frac{\alpha}{2} \sin \frac{\alpha - \alpha^*}{2} \end{aligned} \quad (16)$$

As the goal position is at  $\alpha^* = 0$  the equation simplifies and the angular difference can be approximated by:

$$-2 \sin \frac{\alpha}{2} \sin \frac{\alpha - \alpha^*}{2} \approx -\sin \frac{\alpha}{2} \Delta \alpha \quad (17)$$

This approximation is well suitable for the required range of  $\alpha \in ]-\pi, \pi[$ . A classic Taylor series small-angle approximation exceeds a relative error  $\epsilon_R = 1^\circ$  at  $\alpha = 14^\circ$ , whereas this approximation reaches a relative error of  $\epsilon_R = 1^\circ$  only at  $\alpha = \pm 55.6^\circ$ . Now, combining Equations (15) and (17) provides:

$$\Delta h_{11} \approx -\sin \frac{\alpha}{2} \Delta \alpha + \Delta t_1 \frac{n_1}{d} \quad (18)$$

Analogous to Equation (14) and (18) the formulas for all elements of  $\mathbf{h}$  could be derived and the matrix  $\mathbf{F}$  can be composed to:

$$\mathbf{F} = \mathbf{J}(\frac{\alpha}{2}) \quad (19)$$

Finally, with this matrix the relative position  $\Delta\xi$  could be estimated from the homography measurement using e.g. the pseudo inverse of  $\mathbf{F}$  or a linear solver:

$$\Delta\xi = l(\mathbf{F}) \cdot \Delta\mathbf{h}_m \quad (20)$$

### III. CONTROL FOR PLATOONING

The structure of the proposed controller is depicted in Figure 3, including the necessary extensions for platooning.

For control design purposes, the plant's model considered here is composed of the homography dynamics (8) coupled with the ROMO's vehicle dynamics controller (VDC) [15] and PV's motion dynamics  $\mathbf{u}_{PV}$ . As explained in [16], the VDC provides a control service for the 3 DOF horizontal vehicle motion, capable of tracking the kinematic velocity setpoint  $\mathbf{u}^*$  generated by the platoon controller. More specifically, the VDC determines the necessary steering and torque demands for the ROMO's actuators, such that the requested velocity demands  $\mathbf{u}^*$  are imposed as fast and accurately as possible. Note that the ROMO is able to individually steer and accelerate/brake each wheel, which grants a high degree

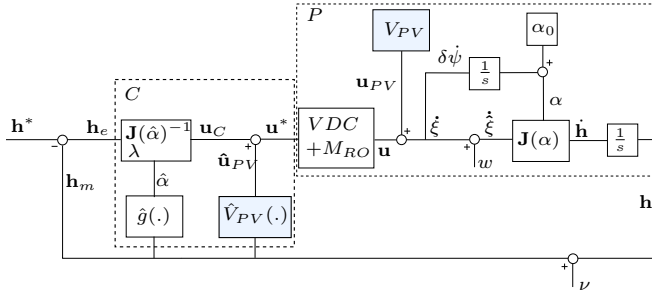


Fig. 3. The Block diagram of the scheme. Light blue blocks are only relevant for the platooning.

of maneuverability. For example: with this configuration the vehicle's lateral and yaw-velocity can be independently controlled, a feature not available in conventional front-steered vehicular configurations, and which is explored in the controller's design.

The proposed platooning controller is composed by two types of control actions:

$$\mathbf{u}^* = \hat{\mathbf{u}}_{PV} + \mathbf{u}_C \quad (21)$$

The first term,  $\hat{\mathbf{u}}_{PV}$ , corresponds to the estimated velocity of the PV and intends to cancel the disturbance effect  $\mathbf{u}_{PV}$  caused by the PV motion in the control loop.

The second term,  $\mathbf{u}_C$ , is a feedback control term that aims at driving the homography  $\mathbf{h}$  to the setpoint  $\mathbf{h}^* = \mathbf{I}$ . The design of these two control actions will be presented in the next two sub-sections.

Before finalizing this initial paragraph, it is useful to introduce two additional exogenous inputs, used in the sequel: i)  $w$  a generic input disturbance, and ii)  $\nu$ , which stands for noise from the sensor and the image processing

#### A. Adaptation for Platooning

Without requiring any further sensor signal the position estimation of the homography method can be utilized to determine  $\mathbf{u}_{PV}$ . Together with a measurement of the own velocity  $\mathbf{u}_m$  and the DHC position estimation (see Equation (20)) an extended Kalman filter (EKF) could be set up to estimate  $\mathbf{u}_{PV}$ :

$$\hat{\mathbf{u}}_{PV} = \hat{V}(\mathbf{h}_{m,t+1}, \mathbf{h}_{m,t}, \hat{\alpha}, \mathbf{u}_m, \Delta t) \quad (22)$$

The advantage of the Kalman filter is the possibility of taking the disturbance  $w$  and the noise  $\nu$  directly into account.

There are different approaches for designing the filter  $\hat{V}$ . In a straight forward approach the filter can estimate  $\hat{\mathbf{u}}_{PV}$  via changes of the relative position. Another approach omits the explicit calculation of the positions and works directly with the relative velocity  $\dot{\xi}$  as the state vector. The basic idea for this is that only the relative velocity over time affects the entries of the homography matrix, so that a detour via the position estimations could be avoided. In this system the state vector is updated by changes of the ROMO's velocity vector  $\Delta \mathbf{u}_t = (\mathbf{u}_t - \mathbf{u}_{t-1})/\Delta t$  and by disturbances  $\mathbf{w}_t$ . Moreover, the measurement model  $M()$  describes the transition of

the system states  $\mu_{t+1}$ , which correspond to the estimated time derivative of the state vector  $\xi$ , to the measurement variable  $z_{t+1}$ , which is equal to the deviation of the stacked homography matrix  $\mathbf{h}_m$  between two time steps. The EKF can be derived from the following set of equations:

$$\begin{aligned} \mu_{t+1} &= \dot{\xi}_{t+1} = \dot{\xi}_t + \Delta \mathbf{u}_t + \mathbf{w}_t \\ z_{t+1} &= \Delta \mathbf{h}_{m,t+1} = M(\mu_{t+1}, \alpha_t, \Delta t) + \nu_t \end{aligned} \quad (23)$$

The measurement formulation is very similar to the position estimation problem from the preceding section. For example,  $\mathbf{F}$  from Equation (13) could be used as the measurement matrix of the EKF's correction step with  $\Delta \xi = \mu \cdot \Delta t$ .

#### B. Controller Design

This section presents the design of the feedback controller  $u_C$ . The design task is carried out in several steps. First, we use again the reduced homography vector  $\underline{\mathbf{h}}$  as in the  $\alpha$  estimation subsection. From the set described in (9) only two are feasible [12] for the platooning case with  $\mathbf{n} = (0 \ 0 \ 1)$ . Here, we select one feasible set (constructed from the last three elements of (9)) to further reduce  $\underline{\mathbf{h}}$  to a 3-dimensional vector  $\underline{\mathbf{h}} = (h_{13} \ h_{31} \ h_{33})^T$ . After reducing the complexity of the model, it is worthwhile now to recall the controller's goal. In particular, we are interested in finding a control law such that  $\underline{\mathbf{h}}$  follows a pre-defined set-point  $\underline{\mathbf{h}}^*$  derived from  $\mathbf{H}^* = \mathbf{I}$ , while fulfilling some transient and robustness requirements (to be specified shortly). With this goal in mind, it is convenient to introduce the homography tracking error  $\underline{\mathbf{h}}_e = \underline{\mathbf{h}}^* - \underline{\mathbf{h}}$  and its dynamics (assuming a constant  $\underline{\mathbf{h}}^*$ ):

$$\dot{\underline{\mathbf{h}}}_e = -\underline{\mathbf{J}}(\alpha)(\mathbf{u} - \mathbf{u}_{PV}) \quad (24)$$

Additionally, it is also important to consider, at the controller design stage, the non-negligible effects of the vehicle controller's inner loops, such as the dynamics of the longitudinal and yaw-rate speed controllers [16]. The dominant dynamics of these inner loops are assumed to be characterized by the following linear representation

$$\dot{\mathbf{u}} = \mathbf{A}_V \mathbf{u} + \mathbf{B}_V \mathbf{u}^* \quad (25)$$

where  $\mathbf{u}^*$  represents the speed setpoints, while the pair  $(\mathbf{A}_V, \mathbf{B}_V)$  characterize the inner loops dynamics. Combining the homography and the actuator models, (10) and (25), renders the following augmented representation:

$$\dot{\mathbf{z}} = \underbrace{\begin{pmatrix} 0 & -\underline{\mathbf{J}}(\alpha) \\ 0 & \mathbf{A}_V \end{pmatrix}}_{\mathbf{A}_z(\alpha)} \mathbf{z} + \underbrace{\begin{pmatrix} 0 \\ \mathbf{B}_V \end{pmatrix}}_{\mathbf{B}_z} \mathbf{u}^* + \begin{pmatrix} \underline{\mathbf{J}}(\alpha) \\ 0 \end{pmatrix} \mathbf{u}_{PV} \quad (26)$$

where  $\mathbf{z} = (\underline{\mathbf{h}}_e \ \mathbf{u})^T$  is the extended state vector. In order to stabilize the dynamics of this augmented system, the following controller law is proposed:

$$\begin{aligned} \mathbf{u}^* &= \hat{\mathbf{u}}_{PV} + \mathbf{u}_C \\ &= \hat{\mathbf{u}}_{PV} + \underline{\mathbf{J}}(\hat{\alpha})^{-1} \lambda \mathbf{E} \mathbf{z} \end{aligned} \quad (27)$$

where  $\lambda$  is a diagonal matrix containing the channel gains  $\lambda_i, i \in \{1, 2, 3\}$ ,  $\mathbf{E} = (\mathbf{I} \ 0)^T$  and  $\hat{\alpha}$  the estimated angle, generated by (12), and represented as:

$$\hat{\alpha} = \alpha + \Delta\alpha \quad (28)$$

with  $\Delta\alpha$  a measurement offset/error. The first component  $\hat{\mathbf{u}}_{PV}$  tries to cancel the effect of the  $\mathbf{u}_{PV}$  disturbance (see previous section for details). The second component relies on the full-state linearization technique [17] and attempts to cancel the effects of the non-linear terms  $\mathbf{J}(\alpha)$ . Under ideal conditions it can be easily verified that the feedback gains  $\lambda_i$  allow us to directly specify the eigenvalues of the homography error dynamics:

**Remark.** If  $\hat{\mathbf{u}}_{PV} \approx \mathbf{u}_{PV}$ ,  $\Delta\alpha \approx 0$  and the inner loops dynamics are neglected, then the homography error dynamics are given by  $\dot{\mathbf{h}}_e = -\lambda \mathbf{h}_e$ . Under these conditions, any  $\lambda_i > 0$  would be enough to stabilize the error dynamics.

However, in practice, the above mentioned assumptions are not always fulfilled, which calls for the robust selection of  $\lambda_i$ , to be carried out in the next sub-section.

### C. Controller's Robust Tuning

To assist in the selection of the controller's gains, (26) and (27) are combined, rendering the following closed-loop dynamics:

$$\dot{\mathbf{z}} = \underbrace{(\mathbf{A}_z(\alpha) - \mathbf{B}_z \mathbf{J}(\hat{\alpha})^{-1} \lambda \mathbf{E})}_{\mathbf{A}(\alpha, \hat{\alpha}, \lambda)} \mathbf{z} + \mathbf{w}_{PV} \quad (29)$$

where  $\mathbf{w}_{PV}$  is regarded as an additive disturbance due to the non-exact cancellation of  $\mathbf{u}_{PV}$ . Additionally, the controller's gains are parametrized as  $\lambda_1 = \lambda_3 = \lambda_P$  and  $\lambda_2 = \iota \cdot \lambda_P$ ; here  $\lambda_P$  is a feedback associated with the translation motion and  $\iota$  is a tuning parameter expressing the ratio between translational and rotational gains. An indicator for choosing  $\iota$  can be the ratio of the time-constants associated with the inner-control loops of the vehicle controller

For tuning and robustness analysis the characteristic polynomial associated with  $\mathbf{A}(\alpha, \hat{\alpha}, \lambda)$  is evaluated using the parameter space approach [18]. The goal here is to determine the set of admissible  $(\lambda_P, \Delta\alpha)$  such that the poles of (29) lie in a pre-defined pole-region (delimited by the so-called  $\Gamma$ -boundary). This specification is also known as  $\Gamma$ -stability in the parameter space literature. Figure 4a depicts the  $\Gamma$ -boundary employed in this work. It is characterized by a shifted imaginary axis  $s = -0.5$ , two lines of constant damping ( $= 1/\sqrt{2}$ ) and a circle centered at the origin and with a radius of 12 rad/s. This boundary is then projected into the parameter space  $(\lambda_P, \Delta\alpha)$  using the methodology described in [18].

In order to also evaluate the dependency of the angular difference  $\alpha$ , the described procedure is carried out for  $\alpha_{-60} = -60^\circ$  (cyan),  $\alpha_0 = 0^\circ$  (grey) and  $\alpha_{60} = 60^\circ$  (magenta). The results (see Figure 4b) are all combined in one parameter space plot, where the respective  $\Gamma$ -stable regions are colored differently. The overlapping area now

provides  $\Gamma$ -stability for the entire range of  $\alpha$  and therefore  $\lambda_P = 1.15$  is chosen to cover the widest expansion of the measurement error  $\Delta\alpha$ . It must be mentioned that even if a set point does not lie in the  $\Gamma$ -stable region, the system can still be stable.

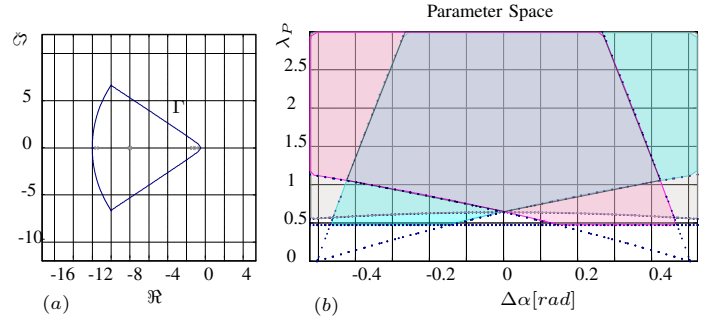


Fig. 4. (a) Specifications of the desired closed-loop pole region ( $\Gamma$ -boundary). (b) The parameter space results with the calculated boundaries and the respective  $\Gamma$ -stable regions for  $\alpha_{-60}$  (cyan),  $\alpha_0$  (grey) and  $\alpha_{60}$  (magenta).

## IV. RESULTS

### A. Position Estimation

In this subsection the performance of the DHC position estimation will be evaluated against the standard homography decomposition method. Two simulation environments will be employed. In the first simulation a camera projection model is utilized to generate feature points that are utilized to calculate the homography matrix. For the second test a 3D environment is simulated, from which artificial images are generated and standard image processing techniques are applied to compute  $\mathbf{H}$ .

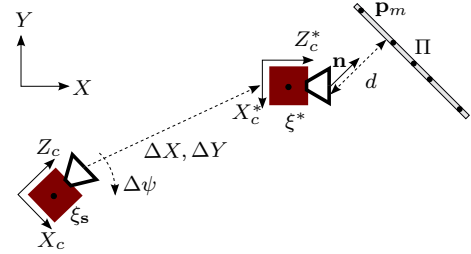


Fig. 5. Simulated scenario for the position estimation from projection model data.

*1) Position Estimation using a Projection Model:* First, a Matlab simulation is set up, where a camera  $C$  moves from a start position  $\xi_s$  to a goal position  $\xi^*$  - see Figure 5. The distance from the initial position to the goal pose is  $\Delta X = 12[m]$ ,  $\Delta Y = 12[m]$  and  $\Delta\psi = 45^\circ$ , while the unit normal vector is  $\mathbf{n} = [-\sqrt{0.5}, 0, \sqrt{0.5}]^T$  and the distance to the feature plane adds up to  $d = \sqrt{18}[m]$ .

Along one run, 48 measurements are taken, where the 3D feature points, all located on the imaginary plane  $\Pi$ , are projected into 2D image coordinates via a camera projection model.



In order to test the algorithm's robustness, normally distributed noise with zero mean and the variance  $\sigma^2$  is added to each 2D feature point. Hence, the true position is not directly derivable from  $\hat{\mathbf{H}}$ . In the following three methods will be compared: i) the homography decomposition [5], ii) the DHC position estimation without determining  $\alpha$  *a priori* (DHC-PE) and iii) the DHC position estimation with the  $\alpha$  determination (DHC-PE $\alpha$ ). These methods receive the noisy homography matrix and then their performance is evaluated using the respective root mean square (RMS) errors for  $\Delta X$ ,  $\Delta Y$  and  $\Delta\psi$ . The ambiguity of several results from the decomposition is omitted by choosing automatically the best solution. During the simulation the motion from  $\xi_s$  to  $\xi^*$  is repeated 500 times for noise with the same error distribution  $\nu_s$ , so that one RMS error value  $\epsilon_{l,rms}$ ,  $l \in \{X, Y, \psi\}$  is calculated from  $48 \times 500$  measurements. The standard deviation  $\sigma_\nu$  is continuously varied from  $[0, 2.5][pix]$  in order to test the behavior under different noise levels. For a better visualization the corresponding error evolutions for  $\Delta X$ ,  $\Delta Y$  and  $\Delta\psi$  are plotted in Figure 6.

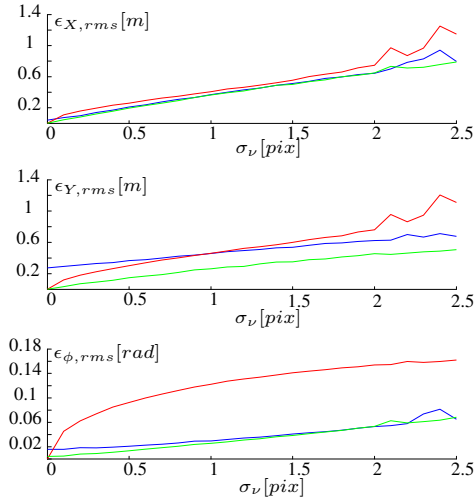


Fig. 6. RMS Errors for  $\Delta X$ ,  $\Delta Y$  and  $\Delta\psi$  for different standard deviations of the noise distribution  $\sigma_\nu$ . The red line denotes the decomposition, the blue line is the DHC-PE and the green graph results from the DHC-PE $\alpha$ .

Along the  $X$  direction, equivalent to the camera's  $Z_C$  direction, the robustness is as expected the lowest due to the physical principle of the sensor. Nevertheless, the proposed DHC method (blue line) performs continuously better than decomposition and the  $\alpha$  extension (green line) improves the RMS error even more. Although the graph of the decomposition error is very similar for both transitional error, the behavior of the DHC methods for the  $Y$  direction is very different to the  $X$  case. For low values of  $\sigma_\nu$  the DHC estimation without the  $\alpha$  extension is slightly more sensitive than the decomposition, but the  $\alpha$  extension decreases the RMS error and the DHC method performs significantly better. It should be remarked that, due to the configuration of  $\mathbf{n}$ , all four analytical solutions are valid and usable. The DHC benefits from this feature, especially for estimating the rotational deviation, which is demonstrated by the low RMS error graphs for  $\Delta\psi$ .

2) *Position Estimation using Artificial Images:* For the second experiment Dymola and the DLR's Visualization Library are utilized to create a simulated 3D environment in combination with physical models. An image of the target plane  $\Pi$ , which is the rear of a vehicle, is taken by the virtual camera at the goal position  $\xi^*$  and afterward the camera is moved to  $q = 5$  different positions  $\xi_j$ , from which the target plane is still visible. The main advantage of this environment is that the ground truth is perfectly known, but the image processing can remain unchanged to the real image case to verify the general functionality. For all images  $I_q$  the  $\hat{\mathbf{H}}$  matrix is calculated using functions from the open source image processing library OpenCV. The feature matcher used for this test is SIFT, which is scale and rotation-invariant to a certain extend [19]. No artificial noise is added to the homography matrices or to the images, as the feature point detection and the matching itself are imperfect due to weak textures, numerical errors etc. and as result the Homography matrix is noisy.

The position estimation scenario is depicted in Figure 7, where part (a) shows the goal view of the camera, in which the distance to the back of the preceding vehicle  $PV$  is  $12.8[m]$  along the camera's optical axis  $Z_C$  respectively the vehicle's  $X$  direction. Except of the back of the van, which is the only relevant section, all other parts of the image are cut out in order to be able to follow the preceding vehicle independently of the environment. An additional texture is placed on the rear of the van, as the original model is poorly textured and features are hard to detect. The feature detection and the generation of the homography matrix are beyond the scope of this work.

Furthermore, Figure 7b shows the camera image at position  $\Delta\xi_5$ , specified in Table I. Each colored line connects a feature points  $\mathbf{p}$  found in the target image  $I^*$  with the equivalent point in the current image  $I$ . Additionally, a green rectangle is drawn, which denotes the detected  $I^*$  in the  $I$ .

The image processing provides the homography matrix  $\bar{\mathbf{H}}$ ,

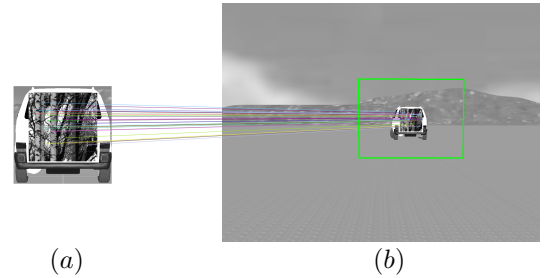


Fig. 7. Features are mapped from the target view (a) to the current view (b) to calculate  $\hat{\mathbf{H}}$ . Additionally, the green frame in (b) visualizes the estimated target image plane  $I^*$ .

from which first the camera parameters contained in the calibration matrix  $\mathbf{K}$  have to be eliminated:

$$\hat{\mathbf{H}} = \mathbf{K}^{-1} \bar{\mathbf{H}} \mathbf{K}$$

Now, the positions are estimated by the DHC position estimation with  $\alpha$  extension and by the homography decom-

position. The right solution of the decomposition is chosen by comparing the normal vector with two feasible estimated solutions. During the test the camera is moved to different

	$\Delta\xi_1$	$\Delta\xi_2$	$\Delta\xi_3$	$\Delta\xi_4$	$\Delta\xi_5$
$\Delta X[m]$	-2	2	3	11	24
$\Delta Y[m]$	0	0	0	2	2
$\Delta\alpha[^\circ]$	0	0	-15	11	0

TABLE I  
RELATIVE POSITIONS  $\xi_i$  TO THE TARGET POSE

positions behind the preceding vehicle, whereas the deviation vectors  $\Delta\xi_j, j \in 1, \dots, q$  for  $q = 5$  positions are given in Table I. The corresponding errors  $\epsilon_i, i \in \{X, Y, \psi\}$  are depicted Figure 8 as a ratio to a norm-error  $\kappa$ . This norm-error is introduced to be better able to compare the errors between both methods and at different positions, whereas the norm-error is chosen such that  $\kappa_X = 0.1[m]$ ,  $\kappa_Y = 0.2[m]$  and  $\kappa_\psi = 1[^\circ]$ .

For closer distances, like  $\Delta\xi_{1,2,3}$ , the error is very small

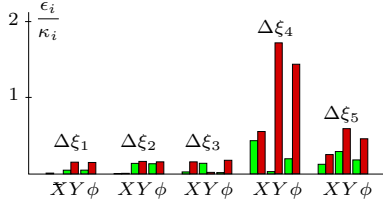


Fig. 8. The ratio between the errors  $\epsilon_i$  and a norm-error  $\kappa_i$  for all three states at 5 different test positions  $\xi_j$ . The green bar expresses the value of the DHC-PE $\alpha$  and the red bar stands for the decomposition error.

with a slightly better performance of the DHC estimation. As expected the errors increase with the distance, whereas the large errors at  $\Delta\xi_4$  result out of the combination of a larger distance with a not negligible lateral and orientational deviation. The DHC-PE $\alpha$  handles the more challenging cases  $\Delta\xi_{4,5}$  clearly better than the decomposition.

The sensitivity of the lateral direction is often significantly larger than of the longitudinal direction, although  $\kappa_X$  is chosen larger than  $\kappa_Y$ . This seems counter-intuitive to the sensor's physical characteristics, but the reason is the way of calculating the homography matrix, which is derived from feature points that are motion-invariant only to a certain extend. Therefore, positions containing larger angular deviations from the goal position are problematic, too. A solution might be the use of a different approach like the homography calculation from template tracking [10].

### B. Platooning

The platooning controller depicted in Figure 3 is tested in a simulated platooning scenario. The ROMO's simulation model contains all the relevant parts including the hierarchical control structure, the kinematics, which are described by an adapted double track model, and the sensor and control buses. During the simulation a conventional vehicle drives in front of the ROMO, which is following autonomously with the motion demand coming only from the DHC platooning

module. The only available measurement of the target vehicle/target pose is a homography matrix  $\tilde{\mathbf{H}}$ , which is disturbed by noise and a sensor delay  $t_d = 0.1[s]$ . The target safety distance  $d = 20[m]$  to the preceding vehicle (PV) is known as well as the normal vector  $\mathbf{n} = [0, 0, 1]^T$ .

At the beginning of the simulation the PV drives straight

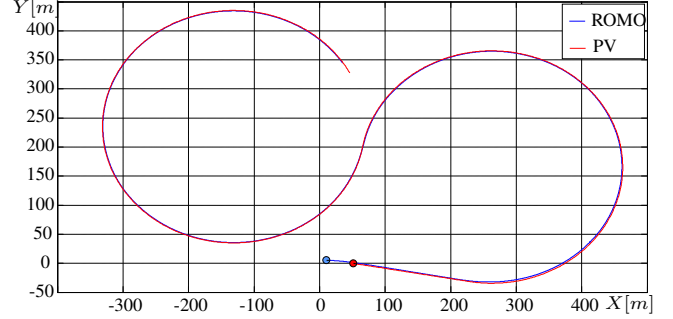


Fig. 9. Motion of the ROMO and PV in the X/Y plane during the simulated run.

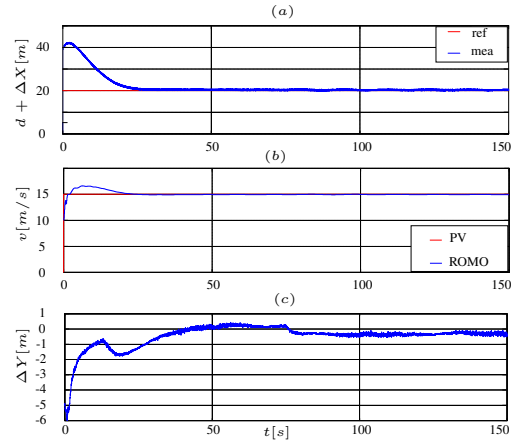


Fig. 10. (a) measured longitudinal gap and the reference distance  $d$ , (b) ROMO's and PV's velocity along the trajectory, (c) lateral gap deviation  $\Delta Y$ .

forward with  $v_{X,PV} = 15[m/s]$ , while the ROMO moves slower with only  $v_X = 10[m/s]$ . The position difference between both is  $\Delta X = 40[m]$  in longitudinal direction and  $\Delta Y = -5[m]$  in longitudinal direction together with an orientation deviation of  $\Delta\psi = -10[^\circ]$ . This initial setup is denoted in Figure 9 by the blue (ROMO) and the red circle (PV), while the black stroke in the circle denotes the orientation.

At the simulation time  $t = 12.2[s]$  the preceding vehicle starts to drive a maneuver consisting of two circle's with a radius of 200 meter - see Figure 9. At this time the ROMO drives with a longitudinal velocity greater then the preceding vehicle's in order to decrease the distance to the goal position - see Figure 10a and b. The lateral control error (Figure 10c) is decreasing until the PV starts its maneuver, where  $\Delta Y$  increases again up to  $-1.8[m]$ . As can be seen in Figure 11b,c the ROMO is compensating the lateral error and the orientation error concurrently by controlling  $v_Y$  and  $\dot{\psi}$ .

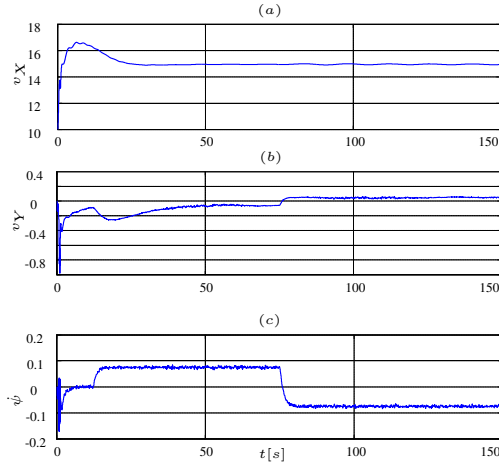


Fig. 11. The ROboMObil's measured motion vector  $\bar{u}$  during the simulation: (a) longitudinal velocity  $v_X$ , (b) lateral velocity  $v_Y$ , (c) yaw rate  $\dot{\psi}$

At around  $t = 25[s]$  the longitudinal target position is reached (Figure 10a) and held by following with PV's velocity (Figure 10b), which is estimated from the homography matrices via the earlier proposed EKF. The lateral gap  $\Delta Y$  is more demanding as the preceding vehicle is constantly yawing, but the absolute lateral error  $\|\Delta Y\|$  is kept below  $0.3[m]$ .

These results demonstrate that the ROMO's simulation model, equipped with the homography-based platoon controller, is able to follow the preceding vehicle within a pre-defined safety distance on a demanding path shaped like an eight. The results also show that the controller is capable of handling considerable initial position and orientation offsets.

## V. CONCLUSION

This paper presented a platooning control scheme based on the direct homography control. By exploiting the homography principle, it was possible to control the 3D motion of the autonomous vehicle while only perceiving 2D images from a monocular camera. Since the controller works exclusively on a space spanned by the homography entries, no decomposition of the homography matrix was necessary. The performance of the DHC was further improved by incorporating an *a priori* estimation of the relative orientation difference increasing the linearization accuracy. This paved the way to the use of DHC for position estimations without requiring the decomposition of the homography matrix. The robustness of these position estimators was demonstrated using images generated from artificial environments. This setting was also employed to verify the performance superiority of the proposed method against the traditional decomposition methods. The DHC scheme was then extended to cope with 2-vehicle platooning scenarios. The robustness of this controller against measurements offsets in the relative orientation was investigated using the parameter space approach. Simulation results, obtained with the ROMO's virtual model, demonstrate a good longitudinal and lateral performance in imitating the preceding vehicle's motion.

As a future work we plan to evaluate the proposed methodology in real world scenarios. Regarding the compensation of the motion of the preceding vehicle the homography measurement could be fused with data from another sensor or source like a Car-2-X module.

## REFERENCES

- [1] Ernst D. Dickmanns. Vehicles capable of dynamic vision. In *IJCAI'97: Proceedings of the Fifteenth international joint conference on Artificial intelligence*, pages 1577–1592, San Francisco, CA, USA, 1997. Morgan Kaufmann Publishers Inc.
- [2] Clemens Rabe, Thomas Müller, Andreas Wedel, and Uwe Franke. Dense, Robust, and Accurate Motion Field Estimation from Stereo Image Sequences in Real-Time. In *Proceedings of the 11th European Conference on Computer Vision*, pages 582–595. Springer, September 2010.
- [3] F. Chaumette and S. Hutchinson. Visual servo control. i. basic approaches. *Robotics Automation Magazine, IEEE*, 13(4):82–90, 2006.
- [4] F. Chaumette and S. Hutchinson. Visual servo control. ii. advanced approaches [tutorial]. *Robotics Automation Magazine, IEEE*, 14(1):109–118, 2007.
- [5] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*. SpringerVerlag, 2003.
- [6] S. Benhimane, E. Malis, P. Rives, and J.R. Azinheira. Vision-based control for car platooning using homography decomposition. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 2161–2166, april 2005.
- [7] A. P. Dani, N. Gans, and W. E. Dixon. Position-based visual servo control of leader-follower formation using image-based relative pose and relative velocity estimation. In *2009 American Control Conference*, pages 5271–5276, June 2009.
- [8] G. Lopez-Nicolas, N.R. Gans, S. Bhattacharya, C. Sagues, J.J. Guerrero, and S. Hutchinson. Homography-based control scheme for mobile robots with nonholonomic and field-of-view constraints. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 40(4):1115–1127, Aug 2010.
- [9] Baoquan Li, Yongchun Fang, and Xuebo Zhang. Projection homography based uncalibrated visual servoing of wheeled mobile robots. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 2167–2172, Dec 2014.
- [10] S. Benhimane and E. Malis. Homography-based 2d visual tracking and servoing. *Int. J. Rob. Res.*, 26(7):661–676, July 2007.
- [11] M.-D. Hua, G. Allibert, S. Krupinski, and T. Hamel. Homography-based visual servoing for autonomous underwater vehicles. In *19th IFAC World Congress*, pages 5729–5733, 2014.
- [12] Alexander Schaub, Tilman Bunte, and Darius Burschka. Direct homography control for highly maneuverable vehicles. In *System Theory, Control and Computing (ICSTCC), 2014 18th International Conference*, pages 826–831, Oct 2014.
- [13] J. Brembeck, L. M. Ho, A. Schaub, C. Satzger, and G. Hirzinger. Romo - the robotic electric vehicle. In *22nd International Symposium on Dynamics of Vehicle on Roads and Tracks. IAVSD*, 2011.
- [14] E. Malis and M. Vargas. Deeper understanding of the homography decomposition for vision-based control. Research Report RR-6303, INRIA, 2007.
- [15] Tilman Bunte, Jonathan Brembeck, and Lok Man Ho. Human machine interface concept for interactive motion control of a highly maneuverable robotic vehicle. In *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pages 1170–1175, june 2011.
- [16] R. de Castro, T. Bunte, and J. Brembeck. Design and validation of the second generation of the robomobil's vehicle dynamics controller. In *International Symposium on Dynamics of Vehicles on Road and Tracks (IAVSD15)*, 2015.
- [17] H.K. Khalil. *Nonlinear Systems*. Pearson Education. Prentice Hall, 2002.
- [18] Jürgen Ackermann. *Robust Control: The Parameter Space Approach (Communications and Control Engineering)*. Springer, 2nd edition, September 2002.
- [19] David G. Lowe. Object recognition from local scale-invariant features. In *Proceedings of the International Conference on Computer Vision - Volume 2 - Volume 2, ICCV '99*, pages 1150–, Washington, DC, USA, 1999. IEEE Computer Society.