

Homography-Based Extrinsic Self-Calibration for Cameras in Automotive Applications

Michael Miksch and Bin Yang
Chair of System Theory and Signal Processing
University of Stuttgart, Germany
{michael.miksch,
bin.yang}@lss.uni-stuttgart.de

Klaus Zimmermann
European Technology Center (EuTEC)
Sony Deutschland GmbH
D-70327 Stuttgart, Germany
klaus.zimmermann@sony.de

Abstract—In this paper we present a method to calibrate the extrinsic parameters of a monocular camera on a moving vehicle. The method is based on a homography which describes the image motion on the road between two camera shots. Therefore, only the road surface has to be visible in the pair of images. A reasonable definition of the world coordinate system in combination with the use of epipolar geometry and odometric data enables the parameterization of the homography matrix with a single parameter. A measure is introduced which evaluates how well the image motion matches to motion described by the homography. The minimization of this one-dimensional error measure finally lead to the extrinsic parameters.

I. INTRODUCTION

Camera calibration is an important area in computer vision. It establishes the relationship between the 3D environment and its projection onto the image plane. The task of calibration can be subdivided into two parts: the intrinsic and the extrinsic calibration. The extrinsic parameters describe the relative position and orientation of the camera with respect to the world coordinate system (WCS). The intrinsic parameters model the projection of points from the camera coordinate system (CCS) onto the image plane.

The parameters are often estimated offline before the initial operation of the system. Methods like [1–4] make use of calibration objects of known geometry. As such objects are not available during the runtime of the system, the parameters are then kept constant. The assumption that the intrinsic parameters are static is reasonable for automotive applications. In contrast, mechanical stress like when the vehicle hits a bump can cause a drift of the extrinsic parameters what degrades the performance of the complete system. Therefore, there is a demand to automatically recalibrate the extrinsic parameters online because a

repetitive calibration with offline methods is not practicable for automotive applications. This paper provides an efficient solution to this problem.

Calibration methods which do not require calibration objects and estimate the parameters during the runtime of the system are called self-calibration. Assumptions on the structure of the road are commonly used as an alternative of calibration objects. [5] and [6] assume straight road boundaries and lane markings, respectively. Dashed and periodic lane markings are used in [7]. Some of the previous principles additionally assume that the vehicle moves in a straight line. These assumptions are often violated under real-world automotive conditions and lead to inaccurate extrinsic parameters. At least for a side view camera, the previous methods are not applicable to solve the calibration problem.

In this paper, only the road surface has to be visible in the images and is assumed to be approximately flat in the immediate vicinity of the camera. The image motion on the road surface, which is induced by the movement of the camera, can be described by a homography. The extrinsic parameters are part of this homography and can be determined if the homography matrix is known. The estimation of the homography matrix is in general difficult because it contains 8 degrees of freedom (DOF). We simplify the model of the homography matrix by first estimating the essential matrix of the epipolar geometry and the vehicle's motion based on odometric data. This enables the introduction of an one-dimensional cost function whose minimization results in the extrinsic parameters.

The outline of this paper is as follows: In Sec. II we introduce the fundamentals. The proposed camera calibration method is described in Sec. III. The results of the method for a real-world video sequence are presented in Sec. IV. Finally, a conclusion is given in Sec. V.

II. BASICS

A. Camera model

The underlying camera model is the pinhole camera, which describes the projection of a 3D object point $\mathbf{M}_w = [X, Y, Z]^T$ onto its image point $\mathbf{m}_p = [u_p, v_p]^T$ by

$$\widetilde{\mathbf{m}}_p \sim \mathbf{A} [\mathbf{R}_e \ \mathbf{t}_e] \widetilde{\mathbf{M}}_w. \quad (1)$$

$\widetilde{\mathbf{M}}_w = [X, Y, Z, 1]^T$ and $\widetilde{\mathbf{m}}_p = [u_p, v_p, 1]^T$ are in homogeneous coordinates, $[\mathbf{R}_e \ \mathbf{t}_e]$ are the extrinsic parameters, and \mathbf{A} is the intrinsic matrix. The homogeneous coordinates are defined up to an arbitrary scaling. Therefore, the operator " \sim " in (1) means

$$\exists \lambda \in \mathbb{R} \setminus \{0\} : [\mathbf{m}_p^T, 1]^T = \lambda \widetilde{\mathbf{m}}_p. \quad (2)$$

The intermediate step

$$\widetilde{\mathbf{m}}_c \sim \mathbf{M}_c = \mathbf{R}_e \mathbf{M}_w + \mathbf{t}_e \quad (3)$$

transforms the point \mathbf{M}_w from the WCS to the point \mathbf{M}_c in the CCS by an Euclidean transform, where \mathbf{R}_e is a rotation matrix with $\mathbf{R}_e^{-1} = \mathbf{R}_e^T$ and \mathbf{t}_e is a translation vector, resulting in normalized coordinates $\widetilde{\mathbf{m}}_c = [u_c, v_c, 1]^T$.

The intrinsic transform between the normalized and the image coordinates is defined by

$$\widetilde{\mathbf{m}}_p \sim \mathbf{A} \widetilde{\mathbf{m}}_c \quad \text{and} \quad \widetilde{\mathbf{m}}_c \sim \mathbf{A}^{-1} \widetilde{\mathbf{m}}_p. \quad (4)$$

B. World coordinate system (WCS)

Without loss of generality, let us consider a WCS whose origin is located on the road surface and below the camera's center of projection (COP). The z -axis of the WCS is pointing towards the camera, whereas the x -axis is pointing parallel to the lateral profile of the vehicle into the direction of travel. As a result, the x - and y -axis describe the road surface which is assumed to be flat.

C. Extrinsic parameters

The extrinsic parameters $[\mathbf{R}_e \ \mathbf{t}_e]$ are already defined in (3). The Euclidean transform can be inverted as follows

$$\mathbf{M}_w = \mathbf{R}_e^T \mathbf{M}_c + \mathbf{t}_h \quad \text{with} \quad \mathbf{t}_h = -\mathbf{R}_e^T \mathbf{t}_e. \quad (5)$$

Based on the assumption in II-B, the COP has the coordinates $\mathbf{M}_c = \mathbf{0}$ in the CCS and $\mathbf{M}_w = \mathbf{t}_h = [0, 0, h_c]^T$ in the WCS where h_c denotes the camera height. Since

$$\mathbf{t}_e = -\mathbf{R}_e \mathbf{t}_h, \quad (6)$$

the extrinsic parameters can be also expressed by the rotation matrix \mathbf{R}_e and the camera height h_c .

D. Motion of the vehicle

The motion of the vehicle between two camera shots is described by an Euclidean transform in the WCS

$$\mathbf{M}'_w = \mathbf{R}_w \mathbf{M}_w + \mathbf{t}_w. \quad (7)$$

The relationship between different coordinates of the same object point in WCS and CCS is illustrated in Fig. 1. The extrinsic parameters are assumed to be identical for both camera shots.

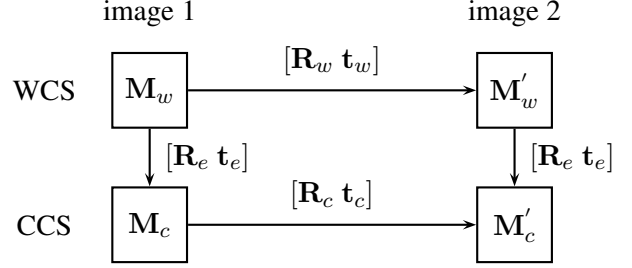


Fig. 1. Relationship between different coordinates

E. Homography of the road surface

The previous definitions lead to the homography matrix between two camera shots for a point on the road surface

$$\mathbf{H}_c = \mathbf{R}_e \left(\mathbf{R}_w + \frac{\mathbf{t}_w \mathbf{n}_w^T}{-h_c} \right) \mathbf{R}_e^T. \quad (8)$$

$\mathbf{n}_w = [0, 0, 1]^T$ is the normal vector of the road surface (see appendix I for a detailed derivation).

III. CALIBRATION

A. Calibration principle

The extrinsic calibration aims at identifying the extrinsic parameters. Besides the vehicle's motion $[\mathbf{R}_w \ \mathbf{t}_w]$, they are the remaining components of the homography matrix in (8). Therefore, the task is to find a homography matrix, which fits best to the real image motion on the road surface between two consecutive frames, in order to get the desired extrinsic parameters.

In Fig. 2, two successive images of a video sequence are shown. The real image motion is visualized by the feature pair (green line) and the red region and its transformed version in the frame below. The motion between the two feature points is defined by

$$\widetilde{\mathbf{m}}'_p \sim \mathbf{A} \mathbf{H}_c \mathbf{A}^{-1} \widetilde{\mathbf{m}}_p. \quad (9)$$

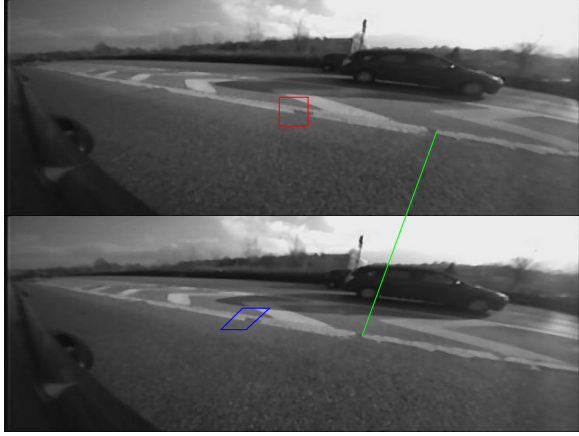


Fig. 2. Two successive frames of a video sequence with one exemplary feature pair (green line) and a rectangular region of size 32×32 (red) and the equivalent transformed region (blue)

The idea is to compare the original and the transformed version of the selected region. For example the sum of absolute differences (SAD) can be computed by

$$e(\mathbf{H}_c) = \sum_{\tilde{\mathbf{m}}_p \in \mathbf{B}} |\mathbf{I}_2(\mathbf{A} \mathbf{H}_c \mathbf{A}^{-1} \tilde{\mathbf{m}}_p) - \mathbf{I}_1(\tilde{\mathbf{m}}_p)| \quad (10)$$

to obtain an error measure, where \mathbf{I}_1 and \mathbf{I}_2 represent both images and \mathbf{B} is the selected region. The homography matches best, if the measure $e(\mathbf{H}_c)$ is minimized.

A homography matrix has 8 DOF because it is defined up to an unknown scaling factor. Hence we need an 8-dimensional search in order to minimize $e(\mathbf{H}_c)$ in (10). To reduce the enormous computational cost, iterative optimization techniques are used. They are applied e.g. in [8] to identify planar regions. They suffer from the fact that an initial guess is needed and the gradient of $e(\mathbf{H}_c)$ has to point towards the global minimum. Unfortunately, the direction of the gradient depends on the selected size and content of the region \mathbf{B} . The larger the region is, the higher the probability is to find the global minimum, but the higher the computational effort is.

To avoid this high-dimensional search, we simplify the parameterization of the homography matrix $\mathbf{H}_c(\alpha)$ to 1 DOF, namely the parameter α . We not only profit from odometric data to determine the vehicle's motion, but also apply the epipolar geometry to achieve this. Finally, the minimization of the one-dimensional error measure $e(\mathbf{H}_c(\alpha))$ leads to the extrinsic parameters. The major steps of this calibration method are summarized in Fig. 3.

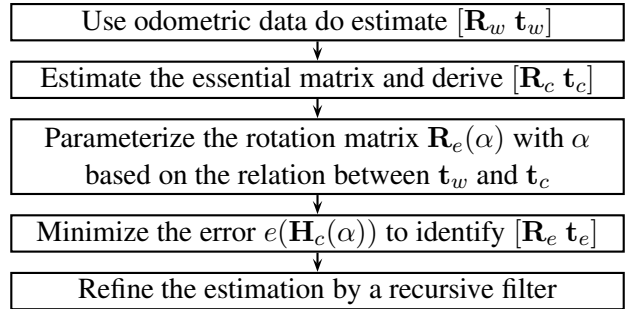


Fig. 3. Overview of the major steps of the calibration method

B. Use of odometric data

Modern vehicles are equipped with sensors, which allow a precise estimation of the vehicle's trajectory [9]. We use the velocity and the yaw rate of the vehicle to estimate a planar trajectory. This means that the rotation is around the z -axis of the WCS (cf. (23)). The motion is defined with respect to the center of gravity (COG), since many sensors are calibrated in reference to the COG. With the knowledge about the position of the camera, it is possible to determine the Euclidean transform $[\mathbf{R}_w \ \mathbf{t}_w]$ between two camera shots, as defined in Sec. II-D. The vehicle moves a certain distance between two camera shots, which is denoted by $\Delta s = \|\mathbf{t}_w\|$ in the following.

C. Estimation of the essential matrix

The estimation of the essential matrix is not covered in this paper, but we refer to the excellent work [10]. It is worth to mention that the estimation of the essential matrix is not restricted to feature pairs on the road. Once the essential matrix is estimated, it can be decomposed into the rotation matrix \mathbf{R}_c and the normalized translation vector $\mathbf{t}_c/\|\mathbf{t}_c\|$ (see [11] for details). The real length of the translation vector can not be derived, because the essential matrix is defined up to an unknown scaling factor.

D. Parameterization of the rotation matrix

The rotation matrix \mathbf{R}_e is normally parameterized by 3 DOF, i.e. one parameter for each rotation angle (e.g. Euler angles). In the following, we reduce the number of required parameters to one, namely the parameter α . According to (21) and (24), $[\mathbf{R}_w \ \mathbf{t}_w]$ and $[\mathbf{R}_c \ \mathbf{t}_c]$ are related by

$$\mathbf{t}_c = \mathbf{R}_e \mathbf{t}_w, \quad (11)$$

$$\mathbf{R}_c = \mathbf{R}_e \mathbf{R}_w \mathbf{R}_e^T. \quad (12)$$

Our basic idea is to exploit the fact that the motion of the camera in the CCS is defined by the translation

vector \mathbf{t}_c and the equivalent translation in the WCS is represented by \mathbf{t}_w . With the use of (11), only one parameter α is enough (see appendix II) to parameterize \mathbf{R}_e . Hence we use the notation $\mathbf{R}_e(\alpha)$ in the following. Equation (12) is implicitly fulfilled with $\mathbf{R}_e(\alpha)$, if the vehicles's motion is consistent with our definitions, and cannot be used as an additional constraint.

E. Proposed calibration method

In summary, two parameters remain in the homography matrix in (8), namely α and h_c . The estimation of both values at the same time by the minimization of the error measure in (10) is challenging, because different combinations of α and h_c result in an almost similar error measure. This fact will be analyzed in detail in Sec. IV-B. Therefore, a constant and known camera height is assumed, which is often a reasonable assumption. Only one parameter is left in the homography matrix

$$\mathbf{H}_c(\alpha) = \mathbf{R}_e(\alpha) \left(\mathbf{R}_w + \theta \frac{\mathbf{t}_w \mathbf{n}_w^T}{\|\mathbf{t}_w\|} \right) \mathbf{R}_e^T(\alpha) \quad (13)$$

where $\theta = \frac{\Delta s}{-h_c}$. That is why we use the notation $\mathbf{H}_c(\alpha)$. Consequently, the rotation matrix is determined by $\mathbf{R}_e(\hat{\alpha})$ with

$$\hat{\alpha} = \operatorname{argmin}_{\alpha \in [-\pi, \pi]} e(\mathbf{H}_c(\alpha)). \quad (14)$$

Finally, the matrix \mathbf{R}_e is converted into the Rodrigues notation [10] because the parameterization $\mathbf{R}_e(\alpha)$ is time varying, just like \mathbf{t}_c and \mathbf{t}_w . The three Rodrigues parameters for any rotation matrix $\mathbf{R}_e = [r_{ij}]_{1 \leq i, j \leq 3}$ are defined as follows

$$\mathbf{w}_{rod} = \frac{v}{2 \sin(v)} [r_{32} - r_{23}, r_{13} - r_{31}, r_{21} - r_{12}]^T \quad (15)$$

with

$$v = \arccos \left(\frac{1}{2} \operatorname{trace}(\mathbf{R}_e) - \frac{1}{2} \right). \quad (16)$$

IV. RESULTS

A. Results for one image pair

For now we consider the images in Fig. 2 again. The important quantities between both camera shots are: $\Delta s \approx 52$ cm, $\Delta \omega \approx 0.02^\circ$ and $h_c \approx 92$ cm. The error is calculated with (10) for the rectangular region marked in red. The error measure $e(\mathbf{H}_c(\alpha))$ for different values of α is depicted in Fig. 4. Its global minimum is around $\alpha \approx -3.0$. The rotation matrix is $\mathbf{R}_e(-3.0)$ and the corresponding Rodrigues parameters are $[1.9185, 0.4581, -0.2130]^T$.

In comparison, the parameters from an offline calibration are $[1.9058, 0.4542, -0.2172]^T$. We see, our self-calibration result is almost similar to the offline calibration result.

An additional fact is visible in the figure. Even though a small range of α is chosen, the gradient of $e(\mathbf{H}_c)$ does not always point to the global minimum. This is a problem, if gradient-based optimization techniques are used and no a priori information is given. If the rotation matrix is approximately known, the range of α can be limited. Otherwise, the minimum should be scanned in the complete range of $\alpha \in [-\pi, \pi]$.

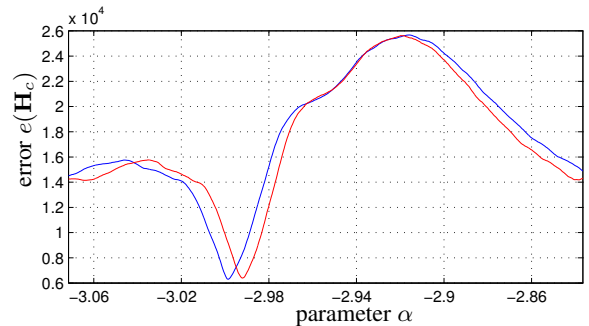


Fig. 4. Error measure for different values of α for a 32×32 region with correct camera height (blue) and wrong camera height (red)

B. Ambiguity for the estimation of α and h_c

In the following, we consider the problem of estimating α and h_c at the same time. It is equivalent to estimate α and θ , since $h_c = \frac{\Delta s}{-\theta}$. The red curve in Fig. 4 shows the error measure if the correct value of θ is scaled by a factor of 1.1. As a consequence, the global minimum is slightly shifted to another position and the extrinsic rotation matrix differs from the correct one. The absolute error measure is still quite similar. For this reason, the joint parameter estimation of α and h_c is challenging because they are estimated based on the minimum of the error measure. The assumption of a static camera height simplifies the estimation, because only α has to be determined and the ambiguity problem is bypassed.

C. Uncertainty of odometric data and camera height

At this point, it is necessary to quantify the rotational error which arises from a wrong camera height or a wrong distance Δs . The uncertainty of θ is modeled by $\theta = \beta \theta_o$ where θ_o is the correct value for θ and β is a multiplicative factor. This leads to an inaccurate estimate of α as follows

$$\hat{\alpha}(\beta) = \operatorname{argmin}_{\alpha \in [-\pi, \pi]} e(\mathbf{H}_c(\alpha, \beta)). \quad (17)$$

The correct rotation matrix $\mathbf{R}_e(-3.0)$ is compared with the estimated matrix $\mathbf{R}_e(\hat{\alpha}(\beta))$ by

$$\mathbf{V}(\beta) = \mathbf{R}_e^T(\hat{\alpha}(\beta)) \mathbf{R}_e(-3.0), \quad (18)$$

$$r(\beta) = \arccos\left(\frac{1}{3}\text{trace}(\mathbf{V}(\beta))\right). \quad (19)$$

The measure $r(\beta)$ approximates the expected value of the angle between $\mathbf{R}_e(\hat{\alpha}(\beta))\mathbf{v}$ and $\mathbf{R}_e(-3.0)\mathbf{v}$ for a random vector \mathbf{v} . In Fig. 5, the error $r(\beta)$ is depicted for different values of β . The accuracy of \mathbf{R}_e is acceptable with respect to a slightly wrong value of θ . For a divergence of 5%, i.e. in our case $h_c \pm 4.6$ cm, the error $r(\beta) \approx 1^\circ$. Even if the car is fully loaded, the height of the camera might not change so much.

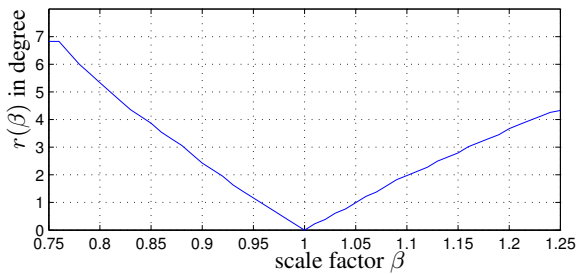


Fig. 5. The rotational error for different values of β , where β is a multiplicative uncertainty factor of θ

D. Results for a video sequence

In the following, a complete video sequence of 60,000 frames is processed to analyze the stability of the proposed calibration method. Actually, the images in Fig. 2 are two frames of this sequence. The camera operates with a resolution of 640x240 pixels and a frame rate of 30 fps. Of course, we need the real extrinsic parameters as a reference for the evaluation of our system. Therefore, the reference parameters are obtained offline with a classical calibration method [1] at the end of the measurement. It uses a checkerboard pattern as calibration object.

The region \mathbf{B} is individually placed for each frame within a region of interest where the road is expected. Additionally, the region \mathbf{B} has a high variance in the intensity values to enable an unambiguous minimization of $e(\mathbf{H}_c)$. In average, every sixth image of the sequence allows a suitable selection of \mathbf{B} . Consequently, only 10,000 estimates of the rotation matrix are taken into account in the histogram plot in Fig. 6. Each color in the histogram represents the distribution of a component in the Rodrigues notation, which has a noticeable peak at a certain value. It is obvious that the method has a low error rate, since most of the estimates are close to that peak.

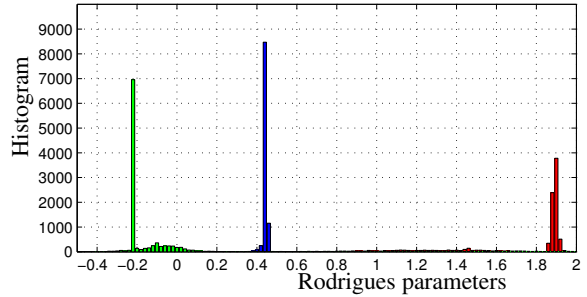


Fig. 6. The histogram over the three Rodrigues parameters of the rotation matrix, a total number of 10,000 estimates

The final estimated Rodrigues parameters and the parameters from the offline calibration are presented in Table I. The error is $\arccos\left(\frac{1}{3}\text{trace}(\mathbf{R}_{\text{ref}}^T \mathbf{R}_e)\right) \approx 0.35^\circ$ between the final extrinsic rotation matrix \mathbf{R}_e and the reference matrix \mathbf{R}_{ref} .

parameters of \mathbf{w}_{rod}	(red)	(blue)	(green)
reference (offline)	1.9058	0.4542	-0.2172
estimated (online)	1.9057	0.4584	-0.2094

TABLE I

ESTIMATED PARAMETERS VERSUS THE REFERENCE VALUES

The final parameters are obtained by a recursive filter, which approximates an average filter with the aim to locate the peak in the distribution. Only those measurements are taken into account, which are inside a certain window. The window is centered around the previous estimated value and the width is adjusted automatically, so that approximately 50% of the estimates are taken into account. In this way, wrong estimates are rejected and do not corrupt the final result.

V. CONCLUSION

We present a novel method to automatically calibrate the extrinsic parameters of a camera system online. A general homography matrix contains 8 DOF. With the use of epipolar geometry and odometric data, 2 DOF remain. An error measure based on the homography matrix and the real image motion is introduced. This measure is minimized in order to estimate the extrinsic parameters. The assumption of a static and known camera height is reasonable to reduce the parameter space to 1 DOF and to prevent ambiguities during the minimization of the error measure. Experiments with a real-world video sequence show that our method results in accurate and stable extrinsic parameters which are comparable to classical offline calibration techniques.

APPENDIX I

It is known in literature [10] that a homography matrix for a plane Π_c is defined by

$$\mathbf{H}_c = \mathbf{R}_c + \frac{\mathbf{t}_c \mathbf{n}_c^T}{d_c} \quad (20)$$

where $[\mathbf{R}_c \ \mathbf{t}_c]$ is the Euclidean transform between two camera views and the plane $\Pi_c : \mathbf{n}_c^T \mathbf{M}_c - d_c = 0$ is defined with respect to the first view. If \mathbf{R}_e and \mathbf{t}_e do not change between these two camera views, $[\mathbf{R}_c \ \mathbf{t}_c]$ can be determined by combining (3), (5), and (7):

$$\mathbf{M}'_c = \underbrace{\mathbf{R}_e \mathbf{R}_w \mathbf{R}_e^T}_{\mathbf{R}_c} \mathbf{M}_c + \underbrace{(\mathbf{R}_e \mathbf{t}_w + \mathbf{t}_e - \mathbf{R}_e \mathbf{R}_w \mathbf{R}_e^T \mathbf{t}_e)}_{\mathbf{t}_c}. \quad (21)$$

A plane $\Pi_w : \mathbf{n}_w^T \mathbf{M}_w - d_w = 0$ defined in the WCS is transformed to a plane in the CCS by

$$\Pi_c : \underbrace{\mathbf{n}_w^T \mathbf{R}_e^T}_{\mathbf{n}_c^T} \mathbf{M}_c - \underbrace{(d_w + \mathbf{n}_w^T \mathbf{R}_e^T \mathbf{t}_e)}_{d_c} = 0. \quad (22)$$

With the definition of the WCS from Sec. II-B, the road surface is defined by $\mathbf{n}_w = [0, 0, 1]^T$ and $d_w = 0$. The assumption that the matrix \mathbf{R}_w has the form

$$\mathbf{R}_w = \begin{bmatrix} \cos(\Delta\omega) & \sin(\Delta\omega) & 0 \\ -\sin(\Delta\omega) & \cos(\Delta\omega) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (23)$$

and the definition of the extrinsic parameters in (6) lead to

$$\mathbf{t}_c = \mathbf{R}_e (\mathbf{R}_w \mathbf{t}_h - \mathbf{t}_h + \mathbf{t}_w) = \mathbf{R}_e \mathbf{t}_w, \quad (24)$$

$$d_c = -\mathbf{n}_w^T \mathbf{R}_e^T \mathbf{R}_e \mathbf{t}_h = -h_c. \quad (25)$$

Substituted into (20), we finally obtain the homography matrix in (8).

APPENDIX II

A rotation matrix can be derived from the Rodrigues rotation formula [10]. It describes a rotation by an angle γ about an axis specified by a unit vector $\mathbf{v} = [v_x, v_y, v_z]^T$ and is given by

$$\mathbf{R} = \mathbf{I} + \mathbf{W} \sin(\gamma) + \mathbf{W}^2 (1 - \cos(\gamma)) \quad (26)$$

where \mathbf{I} is the identity matrix and \mathbf{W} an antisymmetric matrix defined by

$$\mathbf{W} = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}. \quad (27)$$

Rotation matrices which are capable to rotate one vector to another, i.e. $\mathbf{y} = \mathbf{R} \mathbf{x}$, can be parameterized with 1 DOF. The rotation axis is defined by

$$\mathbf{v}(\alpha) = \cos(\alpha) \frac{\mathbf{x} + \mathbf{y}}{\|\mathbf{x} + \mathbf{y}\|} + \sin(\alpha) \frac{\mathbf{x} \times \mathbf{y}}{\|\mathbf{x} \times \mathbf{y}\|} \quad (28)$$

where α is the remaining parameter and the operator "×" is the cross product. The corresponding rotation angle is determined by

$$\gamma = \arccos \left(\frac{\mathbf{p}_x^T \mathbf{p}_y}{\|\mathbf{p}_x\| \|\mathbf{p}_y\|} \right) \quad (29)$$

where $\mathbf{p}_x = \mathbf{x} - (\mathbf{v}^T \mathbf{x}) \mathbf{v}$ and $\mathbf{p}_y = \mathbf{y} - (\mathbf{v}^T \mathbf{y}) \mathbf{v}$ are projections of the vector \mathbf{x} and \mathbf{y} onto the plane, whose normal vector is the rotation axis.

REFERENCES

- [1] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–1334, 1998.
- [2] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3, no. 4, pp. 323–344, 1987.
- [3] S. Hold, C. Nunn, A. Kummert, and S. Müller-Schneiders, "Efficient and robust extrinsic camera calibration procedure for lane departure warning," in *IEEE Proc. Intelligent Vehicles Symposium*, 2009, pp. 382–387.
- [4] M. Bellino, Y. Kolski, and S. Jacot, "Calibration of an embedded camera for driver-assistant systems," in *IEEE Proc. International Conference on Intelligent Transportation Systems*, 2005, pp. 354–359.
- [5] H.-J. Lee and C.-T. Deng, "Camera models determination using multiple frames," in *IEEE Proc. Computer Society Conference on Computer Vision and Pattern Recognition*, 1991, pp. 127–132.
- [6] M. Wu and X. An, "An automatic extrinsic parameter calibration method for camera-on-vehicle on structured road," in *IEEE Proc. International Conference on Vehicular Electronics and Safety*, 2007, pp. 1–5.
- [7] S. Hold, S. Görmer, A. Kummert, M. Meuter, and S. Müller-Schneiders, "A novel approach for the online initial calibration of extrinsic parameters for a car-mounted camera," in *IEEE Proc. International Conference on Intelligent Transportation Systems*, 2009, pp. 420–425.
- [8] M. Okutomi, N. Katsuyuki, M. Junichi, and H. Tomoaki, "Robust estimation of planar regions for visual navigation using sequential stereo images," in *IEEE Proc. International Conference on Robotics and Automation*, 2002, pp. 3321–3327.
- [9] M. Woock, F. Pagel, M. Grinberg, and D. Willersinn, "Odometry-based structure from motion," in *IEEE Proc. Intelligent Vehicles Symposium*, 2007, pp. 1112–1117.
- [10] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.
- [11] R. I. Hartley, "Estimation of relative camera positions for uncalibrated cameras," in *Proc. of the European Conference on Computer Vision*. Springer-Verlag, 1992, pp. 579–587.