

Relative Motion Estimation Based on Sensor Eigenfusion Using a Stereoscopic Vision System and Adaptive Statistical Filtering

Gennaro Notomista, Università degli Studi di Napoli, gennaro.notomista@iit.it, Italy

Alexander Kammenhuber, Audi AG, alexander1.kammenhuber@audi.de, Germany

Parthasarathy Nadarajan, Technische Hochschule Ingolstadt, Parthasarathy.Nadarajan@thi.de, Germany

Michael Botsch, Technische Hochschule Ingolstadt, Michael.Botsch@thi.de, Germany

Mario Selvaggio, Università degli Studi di Napoli, mario.selvaggio@iit.it, Italy

Abstract

This paper presents a method to estimate the relative motion between two vehicles with high accuracy. The estimated quantities are intended to be used as a reference system for automotive sensing techniques and online embedded motion-estimation algorithms. We propose the *sensor eigenfusion* which makes use of a stereoscopic vision system mounted on-board of a host vehicle. Highly reliable markers, i. e., QR-codes, mounted on a remote vehicle are used for robust features detection and tracking. In the case of the mentioned camera system, the proposed method uses the 3D reconstruction capabilities of stereoscopic vision and optical flow techniques usually used in monocular vision systems. The measurements are then shaped, smoothed and fused using a Kalman filter. To achieve the required high accuracy the characteristic statistical parameters of the filter are adapted online according to confidence measures which depend both on the 3D reconstruction and on the optical flow analysis.

1 Introduction

The ability of estimating the motion of mobile robots with respect to the environment in which they are moving is of a great importance. For *autonomous* mobile robots, which aim to navigate themselves in an unknown environment, this is a non-optional feature. Even more important is the case in which the robot requires the motion of other objects that are present in the environment. In many cases, for instance, the objects need to be recognized and categorized as potential obstacles. Another case is represented by the situations in which a moving object is considered as a reference for motion planning functions (e. g. target chasing and dynamic manipulation). A confirmation of what has been mentioned is the great success of competitions such as the DARPA Grand and Urban Challenges [1, 2, 3].

A field in which both the above-mentioned cases are relevant is *autonomous driving*. In order to ensure safe applications in future road traffic, one key challenge is the planning of collision-free paths over time. Since the environment of an autonomous vehicle or robot is not static, the motion of other objects needs to be considered. Other vehicles, pedestrians or animals can, in fact, trigger autonomous intervention functions and, in this way, directly influence the trajectory of the vehicle. In order to optimally plan this trajectory, quantitative information about the motion of the remote objects is required. This paper, therefore, is aimed at presenting a method for estimating the relative position and velocity of the objects present in the environment with respect to a host vehicle. This way the situation can be evaluated and an eventual autonomous intervention can be planned enough time in advance. Commonly used ways of interpreting a situation

are summarized in [4].

Different techniques are already used for environment detection and description as described in [5]. They already make use of several sensor technologies involving radar, laser, infrared, ultrasound and video [6]. Modern cameras, in particular, are very powerful sensors because of the huge quantity of information they can deliver and that can be interpreted in different ways [7]. Methods to exploit this information have been developed in stereoscopic vision systems [8] and in monocular systems that make use of *optical flow* techniques [9]. This latter, inspired by biological creatures such as bees [10] and humans [11], is able to provide information about the relative motion between the observer and the objects as shown in [12] and [13]. Different techniques have been devised in order to combine the depth measurements (i. e., measures of position in the 3D world) with the optical flow vectors (namely measures of velocity in the 2-dimensional image space)[14, 15]. Moreover, also the *flow/depth* constraint can be exploited to detect moving objects in a more robust way [16]. Finally, as reported in [17], optical flow techniques are applied to camera images stabilized using a stereo architecture and knowledge on robot kinematics.

In this paper a stereoscopic system has been designed for the real-time computation of the relative dynamics between the host and the remote systems. The combination of the depth information and the sparse optical flow is performed by means of a statistical filter, whose parameters take into account the relationship between the two measurements techniques. Additionally, the optical flow analysis is performed on the images of both cameras and its output is used as part of the measurement vector in the filtering process. Finally, the statistical parameters of the

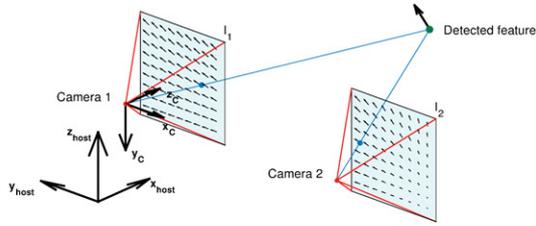


Figure 1: Combination of depth information with optical flow

filter are adapted online based on particular confidence measures related to the measurement process. Therefore, the filter will shape, smooth and merge the data coming from the vision system.

The outline of the paper is as follows. Next section introduces the concept of *eigenfusion* and particularizes it for the presented case. Section 3 formalizes the adaptive statistical filtering. A description of the hardware and the software used in this work is given in Section 4. Section 5 reports the outcomes of the performed tests. Finally, Section 6 concludes the paper.

Throughout this work, vectors and matrices are denoted by lower and upper case bold letters, and random variables are written using sans serif fonts.

2 Sensor Eigenfusion

With the expression *sensor data fusion*, one commonly refers to *multisensor* data fusion, which makes use of measurements coming from different sensor systems in order to get a single estimation of a required quantity, which is most likely to be better than every single measurement [5]. The concept of sensor *eigenfusion*, instead, is based on the assumption that the sensor systems, which provide the measurements, are actually represented by only one sensor type. This way the same information which is contained in the measurements is analyzed using different techniques. Afterwards, the results of these analyses can be merged to obtain a single estimate, which originates then from a deeper exploitation of the information content. This averaging process is weighted, in this case, not on the sensor technology but rather on the technique that is involved in the elaboration of the information. Therefore, the expression of the corresponding weights to be used in the fusion process, does not have a general expression, but it should be derived for the particular case that is considered (see Subsection 2.1).

2.1 Use of the Sensor Eigenfusion in the Presented Case

In the case presented in this paper, the sensor system consists of two cameras and the measurements are therefore the image data. The different techniques involved in the process of information analysis are the 3D reconstruction using the stereoscopic system and the optical flow using

the two monocular systems, of which the stereo-system is made up. In Fig. 1 the described architecture is depicted.

2.1.1 3-D Reconstruction

As can be seen from Fig. 1, the 3-D position of a detected feature is reconstructed by using the properties of the epipolar geometry. Knowing the geometry of the stereo system, represented by the rotation matrix \mathbf{R}_{12} and the translation vector \mathbf{t}_{12} between the two cameras, the following system of equation can be formulated:

$$\begin{cases} \mathbf{p} = \lambda_1 \tilde{\mathbf{p}}_{I1} \\ \mathbf{p} = \mathbf{t}_{12} + \lambda_2 \mathbf{R}_{12} \tilde{\mathbf{p}}_{I2}, \end{cases} \quad (1)$$

where $\mathbf{p} \in \mathbb{R}^3$, is the representation of the point in the reference system of the host vehicle, while $\tilde{\mathbf{p}}_{I1}$ and $\tilde{\mathbf{p}}_{I2}$ are the homogeneous coordinates of a feature in the image planes I_1 and I_2 in Fig. 1. The solution of the system, using least squares techniques, gives the coordinates of the detected marker in the 3-dimensional space.

Moreover, once a feature has been detected in the first image, the epipolar constraint has been exploited to speed up the search for the corresponding feature in the second image [12]. The epipolar constraint is defined by

$$\tilde{\mathbf{p}}_{I1}^T \mathbf{E} \tilde{\mathbf{p}}_{I2} = 0, \quad (2)$$

where the vectors $\tilde{\mathbf{p}}_{I1}$ and $\tilde{\mathbf{p}}_{I2}$ are the two blue rays in Fig. 1, and \mathbf{E} is the essential matrix which is related to the stereo system geometry.

2.1.2 Optical Flow Analysis

The optical flow analysis has already been used in automotive applications and integrated with stereo systems [18, 19, 20, 21, 22, 23, 24, 25]. In the presented paper a *sparse* optical flow technique has been used.

In Fig. 2 two subsequent frames are overlapped: three markers (see Section 4) are depicted in their current and previous configurations, the previous configuration being shadowed. The correspondences between same features detected on the markers is highlighted by red and blue arrows. Starting from these correspondences the velocity vector in the image plane $\mathbf{v}_{Ii} = [\dot{X}_i, \dot{Y}_i]^T$ is computed.

Defining the velocity vector in the camera reference system of Fig. 1 as

$$\mathbf{v} = [{}^C \dot{x}_r, {}^C \dot{y}_r, {}^C \dot{z}_r, {}^C \dot{\varphi}_r, {}^C \dot{\theta}_r, {}^C \dot{\psi}_r]^T, \quad (3)$$

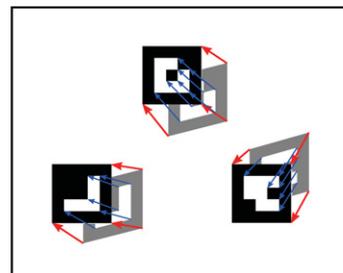


Figure 2: Matched features for the calculation of optical flow vectors

the following equation, relating it to the velocity in the image plane, can be derived:

$$\mathbf{v}_{I_i} = \mathcal{J}_{I_i} \mathbf{v}. \quad (4)$$

The matrix \mathcal{J}_{I_i} is the image Jacobian defined as [12]

$$\mathcal{J}_{I_i} = \begin{bmatrix} -\frac{1}{z_{C_i}} & 0 & \frac{X_i}{z_{C_i}} & X_i Y_i & -(1 + X_i^2) & Y_i \\ 0 & -\frac{1}{z_{C_i}} & \frac{Y_i}{z_{C_i}} & (1 + Y_i^2) & -X_i Y_i & -X_i \end{bmatrix}, \quad (5)$$

where z_{C_i} , the component of the detected feature along the z_C axis, is computed in the camera reference system of Fig. 1.

Combining all the vectors \mathbf{v}_{I_i} , relative to the different detected features, in the vector \mathbf{v}_{I_c} , next equation holds

$$\mathbf{v}_{I_c} = \begin{bmatrix} \mathbf{v}_{I_1} \\ \mathbf{v}_{I_2} \\ \vdots \\ \mathbf{v}_{I_n} \end{bmatrix} = \begin{bmatrix} \mathcal{J}_{I_1} \\ \mathcal{J}_{I_2} \\ \vdots \\ \mathcal{J}_{I_n} \end{bmatrix} \mathbf{v} = \mathcal{J}_{I_c} \mathbf{v}. \quad (6)$$

A least square technique can be employed to solve Eq. (6) for the vector \mathbf{v} . This solution is given by Eq. (7) where $\mathcal{J}_{I_c}^\dagger$ is the image Jacobian pseudo-inverse

$$\mathbf{v} = \mathcal{J}_{I_c}^\dagger \mathbf{v}_{I_c}. \quad (7)$$

The vectors \mathbf{v} are evaluated for the right and the left image and expressed in the same reference system $\mathbf{x}_C \mathbf{y}_C \mathbf{z}_C$ of Fig. 1. The weighted average $\bar{\mathbf{v}}$ of these two vectors ${}^C \mathbf{v}_L$ and ${}^C \mathbf{v}_R$ is defined by

$$\bar{\mathbf{v}} = \frac{\frac{1}{\varepsilon_L} {}^C \mathbf{v}_L + \frac{1}{\varepsilon_R} {}^C \mathbf{v}_R}{\frac{1}{\varepsilon_L} + \frac{1}{\varepsilon_R}}. \quad (8)$$

The expression of the errors ε_L and ε_R is computed using the following definition where also the values of \mathbf{v}_{I_c} , \mathcal{J}_{I_c} and \mathbf{v} are relative to the left and the right image, respectively:

$$\varepsilon = \|\mathbf{v}_{I_c} - \mathcal{J}_{I_c} \mathbf{v}\|. \quad (9)$$

3 Statistical Filtering of Measurements

The measurements provided by the 3D reconstruction and by the optical flow are merged, filtered and fused by the aid of a Kalman filter (KF). This techniques has been applied in [26, 27, 28, 29]. In order to obtain a high accuracy of the estimates, the KF is designed in such a way that it takes account for the nonstationarity of the measurement noise. The noise covariance matrix is chosen adaptively based on confidence measures connected to the 3D reconstruction and the optical flow analysis, as explained in Subsection 3.2.

3.1 System Model

Considering only the planar motion, in Fig. 3 the host and the remote vehicle are depicted in the world reference system $\mathcal{O}XY$.

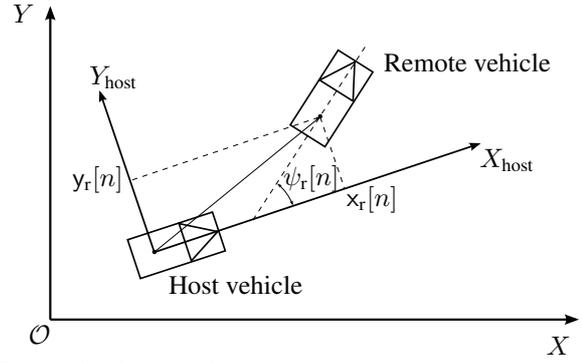


Figure 3: Estimated quantities

The host reference system, whose origin coincides with the mid point of the rear axle of the host vehicle, is also represented. In this reference system, the quantities x_r , y_r and ψ_r , together with their derivatives, are to be estimated. The well-known KF-equations [30] are used for this task. The state and observation equations that describe the dynamic system and which are required to apply the KF, are described in the following.

The state vector $\mathbf{x}[n]$ and measurement vector $\mathbf{y}[n]$ for the KF are introduced as

$$\mathbf{x}[n] = \begin{bmatrix} x_r[n], y_r[n], \psi_r[n], \dot{x}_r[n], \dot{y}_r[n], \dot{\psi}_r[n] \end{bmatrix}^T, \quad (10)$$

$$\mathbf{y}[n] = \begin{bmatrix} x_r^m[n], y_r^m[n], \psi_r^m[n], \dot{x}_r^m[n], \dot{y}_r^m[n], \dot{\psi}_r^m[n] \end{bmatrix}^T, \quad (11)$$

where the superscript ‘‘m’’ indicates that the quantities are measured, i. e., they contain measurement noise. Since the relative dynamics of two objects is to be estimated, a decoupling of $x_r[n]$, $y_r[n]$, and $\psi_r[n]$ is assumed. This way the state equations for x_r and \dot{x}_r are

$$x_r[n+1] = x_r[n] + T\dot{x}_r[n] + \frac{T^2}{2}\eta_{x,s} \quad (12)$$

$$\dot{x}_r[n+1] = \dot{x}_r[n] + T\eta_{x,s}, \quad (13)$$

where T is the time step, and the second derivative of x_r is modelled as process noise $\eta_{x,s}$. Then, the noise variance $\sigma_{\eta_{x,s}}^2$, that is required in the process noise covariance matrix in the KF, is determined using the approximation

$$\sigma_{\eta_{x,s}}^2 \approx \frac{1}{9} (\eta_{x,s}^{\max})^2, \quad (14)$$

where $\eta_{x,s}^{\max}$ is the maximum value of the second derivative of x_r . This approximation is based on the 3σ -rule for Gaussian distributions. The state equations for y_r and \dot{y}_r as well as for ψ_r and $\dot{\psi}_r$ are equivalent to Eqs. (12) and (13). Also the process noise variances for these quantities are approximated equivalently to Eq. (14).

All quantities of the state vector can be measured with the proposed system and so the observation matrix is a simple identity matrix. However, the measurement noise vector $\boldsymbol{\eta}_m \in \mathbb{R}^6$ in the observation equation cannot be assumed to be a stationary random process, for example because of the pixel-quantization in the cameras which leads to a larger noise power for objects that are far away. That is why the measurement noise covariance matrix \mathbf{C}_{η_m} is chosen adaptively in this work.

3.2 Adaptive Measurement Noise Covariance C_{η_m}

Given K measurements $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_K\}$ and the corresponding reference values $\{\mathbf{y}_{\text{ref},1}, \mathbf{y}_{\text{ref},2}, \dots, \mathbf{y}_{\text{ref},K}\}$, which are measured using a device with a very high accuracy, the covariance matrix of the zero-mean noise $\eta_{m,k}$ can be approximated by

$$\tilde{C}_{\eta_m} = \frac{1}{K-1} \sum_{k=1}^K (\mathbf{y}_k - \mathbf{y}_{\text{ref},k}) (\mathbf{y}_k - \mathbf{y}_{\text{ref},k})^T, \quad (15)$$

if the noise process is assumed to be stationary. Since this does not hold for the setting in the current work, it is assumed that the measurement noise process has the property of weak stationarity in certain intervals. These intervals can be specified by the parameters $\varepsilon_L, \varepsilon_R$ defined in Eq. (9) and d , the measure of the distance between the host and the remote vehicle. To implement this approach, the distance d as well as

$$\varepsilon_{LR} = \varepsilon_L + \varepsilon_R \quad (16)$$

are quantized in Q segments, leading to Q^2 intervals where the measurement noise is assumed to be weak stationary. This leads to Q^2 covariance matrices that are estimated using Eq. (15) by using only those data points for their computation which lie in the corresponding intervals. Then, during online operation the task is to determine to which interval the current measurements belong based on the current values of d and ε_{LR} and then to choose the corresponding pre-computed measurement noise covariance matrix $\tilde{C}_{\eta_m} = f(\varepsilon_L, \varepsilon_R, d)$.

4 Hardware and Software Architecture

To validate the presented method of evaluation of the relative motion between a host vehicle and a remote one, a testing procedure has been devised and several tests have been carried out. This section is meant to present the hardware and the software used to perform the tests, while Section 5 is reporting the test setup and their results.

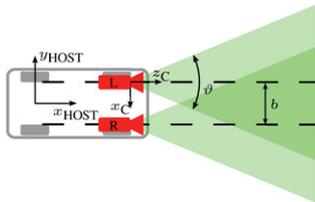


Figure 4: Camera configuration and reference systems

In Fig. 4 the hardware setup is depicted. The two cameras, left and right, are represented together with their horizontal aperture angle ϑ and the baseline of the stereo system b . Furthermore, the reference system of the host vehicle, as well as that of the stereoscopic vision system (which has been chosen to be coincident with that of

the left camera) are introduced. The pattern of the used markers is made up of three highly reliable markers [31] arranged to form a triangle, such that the orientation of the obstacle with respect to the host vehicle can be easily calculated. The 3×3 -bit markers are identified by means of an assigned ID in order to immediately detect their position in the pattern. The entire software has been written in C++ and the developed algorithm has been run on a PC equipped with an Intel Core i7 – 4810MQ and a NVIDIA Quadro K3100M. The used cameras are two mvBlueFOX3 by MATRIX VISION® with a gray scale sensor of 1.3 Mpix (1280×1024), 60 fps and lenses with a horizontal aperture angle θ equal to 40° (see Fig.4).

5 Performance Evaluation



(a) Host vehicle with mounted cameras (left camera) (b) Remote vehicle and DGPS base station

Figure 5: Vehicles setup

In order to validate the developed system, a series of test scenarios have been devised. The tests have been performed on the outdoor test ground of CARISSMA of the Technische Hochschule Ingolstadt. The used vehicles were an Audi A8 of Fig. 5(a) and an Audi S7 of Fig. 5(b) as host and remote vehicle, respectively.

5.1 Referencing Methods

The estimations of the presented system are compared to a ground truth provided by an ADMA (Automotive Dynamic Motion Analyzer) system that makes use of differential GPS (DGPS) corrections. This system is able to provide measurements of position with a precision of less than 2 centimeters at a frequency of 100 Hz. In Fig. 5(b) the base station of the DGPS can be seen. The data were recorded on a PC on board of the host vehicle, to which the cameras were connected via USB and the remote vehicle via Wi-Fi.

5.2 Tests outcomes

The results of the tests are summarized in Fig. 6 for one of the 16 scenarios that have been tested. Here there are 6 plots depicting the time series of the relative longitudinal position (x), lateral position (y), yaw (ψ), longitudinal velocity (\dot{x}), lateral velocity (\dot{y}), yaw rate ($\dot{\psi}$).

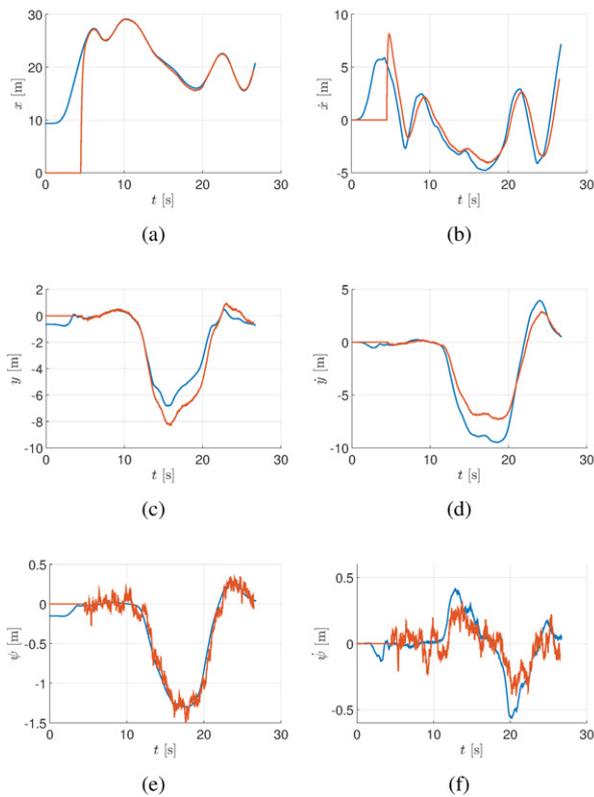


Figure 6: Time series plots relative to a test scenario. On the left column there are: relative longitudinal position (x), relative lateral position (y), relative yaw (ψ); on the right column there are: relative longitudinal velocity (\dot{x}), relative lateral velocity (\dot{y}), relative yaw rate ($\dot{\psi}$).

The ground truth values are plotted in blue, while the estimated quantities are drawn in red. As it can be seen, the tracking capabilities are very good up to high values of relative speed, both for longitudinal and lateral direction. As regards the yaw and the yaw rate estimations, these are more noisy since the yaw is computed and not directly measured, whereas the yaw rate changes much faster than the other quantities in relative motion between vehicles.

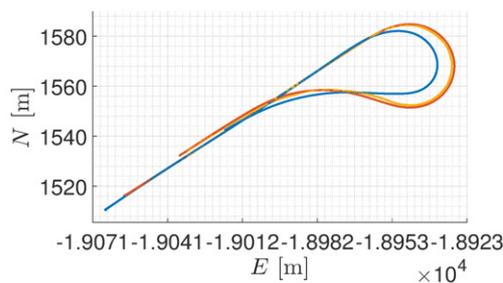


Figure 7: Plot depicting the trajectory of the vehicles while driving a test scenario. Host vehicle trajectory in blue, remote vehicle measured trajectory in red, remote vehicle estimated trajectory in orange.

Fig. 7 represents the same scenario described by the plots in Fig. 6, but in this case the error in space and not in

time is compared. In blue there is the trajectory of the host vehicle, where the cameras were mounted and the data were recorded. The red line is the remote vehicle trajectory measured using the ADMA system, whilst in orange the estimated trajectory is depicted.

6 Conclusion and Future Work

In this paper the potentialities of the sensor *eigenfusion* have been demonstrated for the particular case of video sensors. The devised stereo system is intended to be used as a system for referencing other automotive sensing techniques. In order to improve the estimating capabilities of a statistical filter, particular confidence measures have been defined. These measures are based on the different techniques of analyzing measurement data coming from a single sensor system.

Future work will be focused on refining the relative dynamics estimation by retrieving the kinematic and dynamic model of the remote vehicles, based on previous motion estimations and on object classification. Moreover, a better design of the markers will be devised in order to have faster and more robust estimations. Finally, the developed system will be tested using characteristic features of vehicles, such as the brake lights and the license plate in place of QR-markers, in order to apply it to everyday traffic scenarios.

References

- [1] S. Thrun, “Winning the DARPA Grand Challenge: A robot race through the mojave desert,” in *21st IEEE/ACM International Conference on Automated Software Engineering*, 2006.
- [2] D. Ferguson, M. Darms, C. Urmson, and S. Kolski, “Detection, prediction, and avoidance of dynamic obstacles in urban environments,” in *IEEE Intelligent Vehicles Symposium*, 2008.
- [3] H. Cho, Y.-W. Seo, B. Vijaya Kumar, and R. Rajkumar, “A multi-sensor fusion system for moving object detection and tracking in urban driving environments,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2014.
- [4] S. Lefèvre, D. Vasquez, and C. Laugier, “A survey on motion prediction and risk assessment for intelligent vehicles,” *ROBOMECH Journal*, vol. 1, p. 1, July 2014.
- [5] B. Siciliano and O. Khatib, *Springer Handbook of Robotics*, 1st ed., B. Siciliano and O. Khatib, Eds. Springer-Verlag Berlin Heidelberg, 2008.
- [6] Z. Sun, G. Bebis, and R. Miller, “On-road vehicle detection using optical sensors: A review,” in *Intelligent Transportation Systems Conference*, October 2004, pp. 585 – 590.

- [7] A. Chatterjee, A. Rakshit, and N. N. Singh, *Vision Based Autonomous Robot Navigation. Algorithms and Implementations*, J. Kacprzyk, Ed. Springer-Verlag Berlin Heidelberg, 2013.
- [8] U. Franke and A. Joos, “Real-time stereo vision for urban traffic scene understanding,” in *IEEE Intelligent Vehicles Symposium*, Dearborn (MI), USA, October 2000, pp. 273 – 278.
- [9] H. Chao, Y. Gu, and M. Napolitano, “A survey of optical flow techniques for robotics navigation applications,” *Journal of Intelligent & Robotic Systems*, vol. 73, no. 1-4, pp. 361–372, January 2014.
- [10] N. Franceschini, “Visual guidance based on optic flow: a biorobotic approach,” *Journal of Physiology*, vol. 98, pp. 281 – 292, 2004.
- [11] P. A. Warren and S. K. Rushton, “Optic flow processing for the assessment of object movement during ego movement,” *Current Biology*, vol. 19, no. 18, pp. 1555 – 1560, 2009.
- [12] P. Corke, *Robotics, Vision and Control. Fundamental Algorithms in MATLAB*, B. Siciliano and O. Khatib, Eds. Berlin Heidelberg: Springer-Verlag, 2011.
- [13] R. Szeliski, *Computer Vision. Algorithms and Applications*. Springer London, 2011.
- [14] H. Badino, *Complex Motion*. Springer Berlin Heidelberg, 2007, ch. A Robust Approach for Ego-Motion Estimation Using a Mobile Stereo Platform, pp. 198 – 208.
- [15] U. Franke, C. Rabe, H. Badino, and S. Gehrig, *Pattern Recognition. 27th DAGM Symposium, Vienna, Austria, August 31 - September 2, 2005. Proceedings*. Springer Berlin Heidelberg, 2005, ch. 6D-Vision: Fusion of Stereo and Motion for Robust Environment Perception, pp. 216 – 223.
- [16] S. Heinrich, *Pattern Recognition - 24th DAGM Symposium Zurich, Switzerland, September 16 - 18, 2002 Proceedings*. Springer Berlin Heidelberg, 2002, ch. Real Time Fusion of Motion and Stereo Using Flow/Depth Constraint for Fast Obstacle Detection, pp. 75 – 82.
- [17] J. Schmudderich, V. Willert, J. Eggert, S. Rebhan, C. Goerick, G. Sagerer, and E. Korner, “Estimating object proper motion using optical flow, kinematics, and depth information,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 4, pp. 1139 – 1151, August 2008.
- [18] A. Mitiche and J. K. Aggarwal, *Computer Vision Analysis of Image Motion by Variational Methods*. Switzerland: Springer International Publishing, 2014, ch. Optical Flow Estimation, pp. 41 – 93.
- [19] A. Zisserman and R. Hartley, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.
- [20] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, no. 1-3, pp. 185–203, August 1981.
- [21] D. Patel and S. Upadhyay, “Optical flow measurement using Lucas-Kanade method,” *International Journal of Computer Applications*, vol. 61, no. 10, pp. 6 – 10, January 2013.
- [22] D. Willersinn and W. Enkelmann, “Robust obstacle detection and tracking by motion analysis,” in *IEEE Conference on Intelligent Transportation System (ITSC)*, Boston, MA, 1997, pp. 717 – 722.
- [23] W. Kruger, W. Enkelmann, and S. Rossle, “Real-time estimation and tracking of optical flow vectors for obstacle detection,” in *Proceedings of the Intelligent Vehicles Symposium*. Detroit, MI: IEEE, September 1995, pp. 304 – 309.
- [24] J. T. Philip, B. Samuvel, K. Pradeesh, and N. K. Nimmi, “A comparative study of block matching and optical flow motion estimation algorithms,” in *International Conference on Magnetism, Machines & Drives (AICERA-2014 iCMMD)*, 2014.
- [25] V. Borsu and P. Payeur, “Pose and motion estimation of a moving rigid body with few features,” in *IEEE International Workshop on Robotic and Sensors Environments (ROSE)*, November 2009, pp. 116 – 121.
- [26] L. Tong, *System Simulation and Scientific Computing*. Berlin Heidelberg: Springer Berlin Heidelberg, 2012, ch. An Approach for Vehicle State Estimation Using Extended Kalman Filter, pp. 56 – 63.
- [27] T. Dang, C. Hoffmann, and C. Stiller, “Fusing optical flow and stereo disparity for object tracking,” in *IEEE 5th International Conference on Intelligent Transportation Systems*, Singapore, 2002, pp. 112 – 117.
- [28] S. Gehrig, C. Rabe, and L. Krueger, “6D vision goes fisheye for intersection assistance,” in *Canadian Conference on Computer and Robot Vision (CRV)*. Windsor, Ont.: IEEE, May 2008, pp. 34 – 41.
- [29] J. Eggert, S. Rebhan, J. Schmudderich, and V. Willert, “Estimating objects proper motion using optical flow, kinematics and depth information,” U.S. Patent US 2011/0243 390 A1, October 6, 2011.
- [30] D. Simon, *Optimal State Estimation*. WILEY, 2006.
- [31] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, “Automatic generation and detection of highly reliable fiducial markers under occlusion,” *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, June 2014.