

Accurate Global Localization Using Visual Odometry and Digital Maps on Urban Environments

Ignacio Parra Alonso, David Fernández Llorca, *Member, IEEE*, Miguel Gavilán, Sergio Álvarez Pardo, Miguel Ángel García-Garrido, Ljubo Vlacic, and Miguel Ángel Sotelo, *Member, IEEE*

Abstract—Over the past few years, advanced driver-assistance systems (ADASs) have become a key element in the research and development of intelligent transportation systems (ITSs) and particularly of intelligent vehicles. Many of these systems require accurate global localization information, which has been traditionally performed by the Global Positioning System (GPS), despite its well-known failings, particularly in urban environments. Different solutions have been attempted to bridge the gaps of GPS positioning errors, but they usually require additional expensive sensors. Vision-based algorithms have proved to be capable of tracking the position of a vehicle over long distances using only a sequence of images as input and with no prior knowledge of the environment. This paper describes a full solution to the estimation of the global position of a vehicle in a digital road map by means of visual information alone. Our solution is based on a stereo platform used to estimate the motion trajectory of the ego vehicle and a map-matching algorithm, which will correct the cumulative errors of the vision-based motion information and estimate the global position of the vehicle in a digital road map. We demonstrate our system in large-scale urban experiments reaching high accuracy in the estimation of the global position and allowing for longer GPS blackouts due to both the high accuracy of our visual odometry estimation and the correction of the cumulative error of the map-matching algorithm. Typically, challenging situations in urban environments such as nonstatic objects or illumination exceeding the dynamic range of the cameras are shown and discussed.

Index Terms—Digital maps, geographical information systems, intelligent transportation systems (ITSs), intelligent vehicles, visual odometry.

I. INTRODUCTION

OVER the past few years, advanced driver assistance systems (ADASs) have become a key element in the research and development of intelligent transportation systems (ITS) and particularly of intelligent vehicles. ADASs are becoming in-

Manuscript received July 15, 2011; revised October 26, 2011; accepted March 5, 2012. Date of publication April 30, 2012; date of current version November 27, 2012. This work was supported by the Spanish Ministry of Economy and Competitiveness under Research Grant ONDA-FP TRA2011-27712-C02-02. The Associate Editor for this paper was A. Hegyi.

I. Parra Alonso, D. Fernández Llorca, M. Gavilán, S. Álvarez Pardo, M. Á. García-Garrido, and M. Á. Sotelo are with the Department of Computer Engineering, Polytechnic School, University of Alcalá, 28871 Alcalá de Henares, Spain (e-mail: ignacio@aut.uah.es; david.fernandezl@uah.es; miguel.gavilan@aut.uah.es; sergio.alvarez@aut.uah.es; garrido@depeca.uah.es; miguel.sotelo@uah.es).

L. Vlacic is with the Intelligent Control Systems Laboratory, Institute of Integrated and Intelligent Systems, Griffith University, Brisbane, Qld. 4111, Australia (e-mail: l.vlacic@griffith.edu.au).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TITS.2012.2193569

creasingly more common in modern vehicles, and new sensors come now as standard to improve safety and aid in navigation. These systems usually require precise knowledge of the vehicle's position and orientation, or even its global localization. In addition, research in the field of vehicle-to-vehicle and vehicle-to-infrastructure cooperation for traffic management is now a hot topic in the ITS field. Moreover, many applications require information about not only the local motion of the vehicle but its position in a global map for advance warning of approaching hazards, such as intersections and pedestrian crossings [1] as well. Traditionally, this location information has been provided by Global Positioning System (GPS), despite its well-known limitations, particularly in urban environments (signal blockage, severe multipath, etc.). Different solutions have been proposed to bridge the gap of GPS position errors, but they usually require additional expensive sensors; their drift rapidly increase with time, and frequent calibration is required [2].

Over the past few years, vision-based systems have proved to be capable of locally tracking the position of a vehicle over long distances using only the images as inputs and with no prior knowledge of the environment [3]. Although results are encouraging, there are still some problems to be resolved. First, in dynamic untextured outdoor environments feature points are hard to find and track, resulting in an increase in the number of outliers. Featureless solutions have been proposed based on textures [4] or whole image alignment and near-planarity assumptions [5]. These methods yield good results but have problems with the planar world assumption and with heavy shading from other vehicles or objects surrounding the road due to its usually narrow field of view [5], [6]. They are also limited in the speed of the vehicle due to the camera position, usually facing downward and close to the ground, which will demand high frame rates for high vehicle speeds. Second, the nature of the measurements in vision systems produces higher inaccuracies in the depth direction of the camera [7]. This leads to errors estimating the lengths traveled or the pitch and roll angles, depending on the camera setup and the motion model [8]. Different solutions have been proposed to cope with this problem, being the most common to change the camera configuration to increase the optical flow in the images [6], [9], to estimate the different components of the motion separately using different sets of points and algorithms [8], [9], and to model the uncertainty in the measurements and incorporate it in the optimization problem [10], [11]. Changing the camera's configuration and estimating separately the motion usually come together and deliver good results at the cost of

increasing the complexity of the setup and not allowing for an easy integration of systems. On the other hand, models of the uncertainty are approximations and, thus, only valid under some constraints [12], [13]. Third, a real-world problem that is usually overseen is the disturbances introduced by illumination changes and nonstationary objects. While downward facing systems are more sensitive to illumination changes, nonstationary objects are one of the main challenges for the rest of them [14], [15]. Most of the systems do not deal with this problem or just trust that it will be filtered by a large-enough number of inliers and a good outlier rejection scheme. Finally, a common problem of odometry systems is the drift in the estimation over long distances, which is inherent to the nature of the information available. Visual odometry produces locally accurate and smooth estimates that inevitably drift over time due to the lack of a global reference, which would allow removal of the small cumulative errors made in the successive estimations. The straightforward solution is to fuse this information with a second sensor, providing coarse but absolute position estimations. This way, visual odometry can deliver smooth and locally accurate position estimations, whereas a second sensor will offer coarse but absolute estimations. Examples of this solution have been given using GPS [3], [5], IMUS [16], [17], and GYROS [18].

In this paper, we show that accurate global localization is possible, using a stereo video stream to produce an estimated trajectory that is accurate and robust to illumination changes and nonstationary objects. The motion trajectory of the vehicle is estimated using weighted nonlinear least squares (WNLS) optimization and a Gaussian multivariate model to estimate the uncertainties in the measurements [19]–[21]. The outlier removal is performed by a RANSAC algorithm, based on Mahalanobis distance to better deal with the nature of the input data. In a further step, the local motion information is matched to a digital map providing global positioning and removing the cumulative error. Fusing the locally accurate trajectory information from the visual odometry with a digital road map, we are able to accurately estimate the global position of the vehicle for very long GPS outages, and using map features, the drift errors are corrected. We provide examples of trajectory estimations and map fusion in long urban environments with real traffic and discuss the key issues for further improvement. The main contribution of this work is the demonstration of a full localization solution in real complex environments in the presence of nonstationary objects, under extreme illumination conditions, and for distances on the order of kilometers. This is achieved due to the use of a simplified and calibrated motion model that rejects nonfeasible motions for the vehicle and of features strong to perspective and illumination changes and the fusion with a digital map, which corrects the drift errors inherent to odometry systems.

The remainder of this paper is organized as follows: Section II briefly surveys vision-based egomotion systems. An overall description of the visual odometry system is presented in Section III. The integration and correction of the visual odometry information using GPS and a digital map is described in Section IV. The experimental setup and extensive examples of the system performance are provided in Sections V and VI.

Finally, conclusions and discussion for further improvement can be found in Section VII.

II. RELATED WORK

Early attempts to recover the motion of a camera or cameras using only a video stream came from the robotic community more than two decades ago [22]. The first visual odometry systems were successfully used by the NASA rovers to support their wheel odometry information since early 2004 [23]. However, probably, the first convincing stereo visual odometry system not linked to robotics was presented by [24]. Although monocular systems present good results [25], they still require high-quality features to robustly perform [5] and are more prone to errors [26]. As our work is focused on global localization in complex urban environments, the use of a stereo system is highly recommended to mitigate the disturbances from low-textured environments, nonstationary objects, and changes in illumination.

The idea of estimating displacements from two 3-D frames using stereo vision was previously used in [17], [27], and [28]. A common factor of these works is the use of robust estimation and outlier rejection using RANdom SAMple Consensus (RANSAC) [29]. In most previous research on visual odometry, features are used for establishing correspondences between consecutive frames in a video sequence. Some of the most common choices are the Harris corner detector [30] and the Kanade–Lucas–Tomasi feature tracker [31]. In [32], a new feature extractor was specifically developed for the image registration problem in visual odometry systems, improving both stability and accuracy of previous feature extractors. In this paper, Lowe’s Scale-Invariant Feature Transform (SIFT) [33] was used to cope with the changing illumination and clutter of urban environments at the cost of a higher computational load. At each frame, SIFT features are extracted and matched, and epipolar constraints are applied to reduce the number of bad matches [21].

In [17], a so-called firewall mechanism is implemented to reset the system to remove cumulative error. Another common approach to maintain consistency in long sequences of images is to introduce a *local bundle adjustment* [16], [25]. Although these systems increase the final accuracy in the position estimation, their results are similar to those not using bundle adjustment (about 5% accuracy for [16] and errors on the order of meters for [25]). In the last years, some systems have tried to reduce the drift using some sort of previously known map. The use of previous knowledge of the environment to reduce the drift has been explored in [34], where a map with a simplified 3-D model of the buildings is matched to the measurements along with the camera’s pose in a final global bundle adjustment. In [35], Napier *et al.* used overhead satellite images to reduce the drift of a stereo visual odometry system. In our approach, the topological information provided by a digital map is matched to the motion trajectory of the vehicle, and map features are used to reset the drift of the visual odometry as global localization in the map is performed. To do so, a probabilistic map-matching algorithm constrained to the road matches the vehicle trajectory to global positions in the map.



Fig. 1. Camera coordinate system.

This idea is based on a previous work where GPS and dead reckoning were fused in a similar approach [2].

III. VISUAL ODOMETRY

The problem of estimating the trajectory followed by a moving vehicle can be seen as a continuous optimization problem. The movement between consecutive frames can be uniquely described by a rotation matrix R and a translation vector T that characterize the motion of the scene being observed. The use of nonlinear methods becomes necessary since the rotation matrix is not linear and has to be orthonormal. In addition, the uncertainty in the 3-D position of a reconstructed feature heavily depends on its location. In [36], Aitken showed that, when a weighted sum of squared residuals is minimized, the estimation is the Best Linear Unbiased Estimator (BLUE) if each weight is equal to the reciprocal of the variance of the measurement. In our case, we have used a RANSAC-based WNLS algorithm to minimize the reprojection error. A Gaussian multivariate model of the uncertainty in the 3-D reconstruction of the features is used as weights for the WNLS. In this section, a brief explanation of our visual odometry system is presented. For further details, see [21].

A. Motion Model

In the typical driving scenario, the road forms a planar structure, and the motion of the car can be modeled with three predominant parameters, i.e., forward translation, pitch, and yaw. With this simplification, the motions allowed are reduced, and the estimation is faster and more robust. However, in this model, we assume a coordinate frame in which the ground plane is parallel to the XZ plane of our camera coordinate system and that the optical axis is parallel to the Z -axis (see Fig. 1). This is not true and requires the points to be rectified prior to computing the ego-motion [37].

To estimate the pitch and yaw of the camera's rig with respect to the ideal position, an offline calibration procedure is performed. First, the ego-motion for a video showing a long straight motion on a flat road is estimated by solving the equations only for forward translation and yaw. If the rig has some rotation around the y -axis, the motion model will try to compensate this rotation by bending the trajectory around the

y -axis. We will adjust the yaw value accordingly and repeat until no drift with respect to the straight trajectory is found. We will repeat the experiment for a flat road and the pitch angle.

This calibration values for the pitch and the yaw are then used for the calibration of the cameras. Prior to the ego-motion estimation, the 3-D position of the points will be corrected according to these values to comply with the simplified pitch, yaw, and forward translation model. This approximation, along with the RANSAC outlier rejection step, allows the system to cope with moving objects such as pedestrians or other cars. On the one hand, RANSAC will reject every minimal solution, as long as the number of stationary points being tracked is higher than the outliers (pedestrians or other moving cars). On the other hand, the 2-D approximation adds some information about the car dynamics to the model, rejecting solutions incompatible with the car dynamics and speeding up the algorithm.

B. Optimization

At each frame, SIFT features are extracted from each of the four images (stereo pair at time 0 and stereo pair at time 1) and stereo matched among the stereo pairs. The resulting matches for the stereo pairs are then matched again among them. Only the features finding a matching pair in the three matching processes will be used for the computation of the ego-motion. As the matching of SIFT features does not rely on epipolar geometry, epipolar constraints are applied to reduce the number of bad matches. (For further details, see [21].)

In our optimization problem, the reprojection error of the tracked features will be minimized using a weighted scheme, i.e.,

$$\mathbf{E}_i(\mathbf{x}) = w_i \cdot (\mathbf{f}_i(\mathbf{x}) - b_i)^2, \quad i = x, y, z \quad (1)$$

where i indicates coordinates x , y , or z ; $\mathbf{E}_i : \mathbb{R}^6 \rightarrow \mathbb{R}$ is the reprojection error; w_i are the weights; $\mathbf{f}_i : \mathbb{R}^6 \rightarrow \mathbb{R}$ is the coordinate i reprojection of a 3-D point for the estimated solution $\mathbf{x} = [\theta_x \theta_y \theta_z t_x t_y t_z]$; and b_i is the measured 3-D i coordinate at t_1 . Error functions $\mathbf{E}_i(\mathbf{x})$ are minimized with regard to \mathbf{x} using N pairs of 3-D points, given a value of \mathbf{x} , by means of an iterative process. The minimum value of $\mathbf{E}_i(\mathbf{x})$ occurs when the gradient is zero, i.e.,

$$\frac{\partial \mathbf{E}_i(\mathbf{x})}{\partial \mathbf{x}} = 2w_i \cdot (f_i(\mathbf{x}) - b_i) \cdot \frac{\partial (f_i(\mathbf{x}) - b_i)}{\partial \mathbf{x}} = 0. \quad (2)$$

At each iteration, the model is linearized by approximation to a first-order Taylor's series expansion about \mathbf{x} as given by

$$\frac{\partial \mathbf{E}_i(\mathbf{x})}{\partial \mathbf{x}} \approx 2w_i (f_i(\mathbf{x}) + \nabla f_i(\mathbf{x}) \partial \mathbf{x} - b_i) \frac{\partial f_i(\mathbf{x})}{\partial \mathbf{x}} = 0 \quad (3)$$

where $\nabla f_i(\mathbf{x}) = (\partial f_i / \partial x_1, \dots, \partial f_i / \partial x_6)^t$ is the gradient of f_i calculated at point \mathbf{x} , neglecting high-order terms. Rearranging (3), the normal equations are obtained, which, written in matrix notation, are given by

$$(\mathbf{J}^T \mathbf{W} \mathbf{J}) \delta \mathbf{x} = \mathbf{J}^T \mathbf{W} \mathbf{C} \quad (4)$$

where $\mathbf{W}_{3 \times 3}$ is the weights matrix, $\mathbf{J}_{3 \times 6}$ is the Jacobian matrix, and

$$\mathbf{C}_{3 \times 1} = \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} - \begin{pmatrix} f_x(\mathbf{x}) \\ f_y(\mathbf{x}) \\ f_z(\mathbf{x}) \end{pmatrix}. \quad (5)$$

The system is iteratively solved using

$$\delta \mathbf{x} = (\mathbf{J}^T \mathbf{W} \mathbf{J})^{-1} \mathbf{J}^T \mathbf{W} \mathbf{C} \quad (6)$$

and N pairs of 3-D points until an absolute minimum error is found or a maximum number of iterations is reached.

The covariance estimate of the pose is approximated from the Jacobian \mathbf{J} of the error function as

$$\Sigma_x = (\mathbf{J}^t \mathbf{W} \mathbf{J})^{-1}. \quad (7)$$

IV. ASSISTANCE TO A GLOBAL POSITIONING SYSTEM USING A DIGITAL MAP

Traditionally, GPS and dead reckoning has been used as input to map-matching algorithms; however, the conventional integration does not correct the position using the map information after relocation. Given the cumulative nature of errors in visual odometry estimations, the drift will keep increasing without bounding. Moreover, the complex nature of the urban environment and the numerous nonstatic objects (other cars, pedestrians, etc.) will make the map-matching process unreliable and eventually lose the vehicle position. If accurate localization is needed for long periods of GPS outage, additional information available in the digital map has to be used to correct the vehicle position and reset the cumulative errors from visual odometry. Otherwise, small misestimations due to poor quality of the input images (rain, glares, etc.) or nonstatic objects can quickly lead to mislocalizations.

In our approach, we propose a probabilistic map-matching algorithm constrained to the road, which uses map features to control the errors of the visual odometry by feeding back corrections from the map-matching process. This calibration process looks for sharp turns and roundabouts in the digital map and correct the vehicle position by removing the cumulative error of the visual odometry whenever the vehicle is localized in such map features. Every time the map-matching algorithm correctly matches the vehicle position at one of these map features, the vehicle position and heading are corrected using digital map information about orientation and the length of the stretch that the vehicle is traveling.

A. Integration of Visual Odometry and GPS

The visual odometry motion information is filtered and down-sampled to one position per second to emulate the sampling rate of the GPS. This way, the user will not notice the difference between the GPS and the visual odometry. The position of the vehicle is represented as a distance from a map node and an orientation with respect to the segment between the two nodes that the car is on. The car position is forced to be in the segment between nodes, which is usually known as *lock*

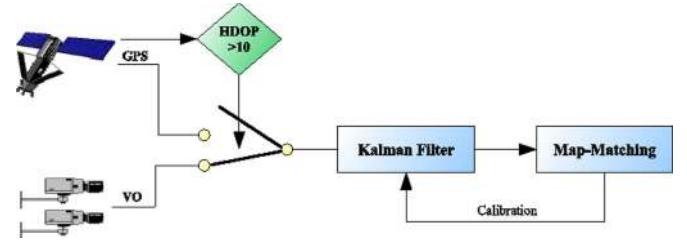


Fig. 2. Integration of the GPS and visual odometry measures.

on road on commercial GPS. As long as the vehicle moves, its position in the map is updated according to the input from the visual odometry system. The signal from each GPS satellite has a level of precision that depends on the relative geometry of the satellites. The quality of the position delivered by the GPS is estimated using the horizontal dilution of position (HDOP). When visible GPS satellites are close together in the sky, the geometry is said to be weak, and the dilution of position (DOP) value is high; when far apart, the geometry is strong, and the DOP value is low.

In our system, when the HDOP is greater than 10, which has been traditionally interpreted as moderate quality, the signal is considered to be not reliable, and the position in the map is computed using the visual odometry information (see Fig. 2).

B. Map Matching

Map-matching algorithms use inputs generated from positioning technologies and supplement this with data from a high-resolution spatial road network map to provide an enhanced positioning output. The general purpose of a map-matching algorithm is to identify the correct road segment on which the vehicle is traveling and to determine the vehicle location on that segment [38], [39]. Map matching not only enables the physical location of the vehicle to be identified but also improves the positioning accuracy if good spatial road network data are available [40].

1) *Identification of the Actual Link*: The most complex element of any map-matching algorithm is to identify the actual link among the candidate links [38]. In our map-matching algorithm, three basic assumptions are made.

- 1) The vehicle travels on the road most of the time.
- 2) The vehicle cannot jump from one place to another one with no connection.
- 3) The vehicle has to follow certain road rules.

First, the initial road segment in which the vehicle is traveling is estimated through an initialization process; when the GPS position is lost, an elliptical confidence region is computed using the visual odometry uncertainty and the last reliable GPS position. The confidence region is projected onto the map, and the road segments that are within the confidence region are taken as candidate regions (see Fig. 3). For simplicity, the elliptical confidence region is approximated to a rectangular one.

If the confidence region contains more than one candidate segment, the heading over the last 5 s is computed and compared with the segment orientation. If there is only one candidate left after the heading check, that is the initial road segment.

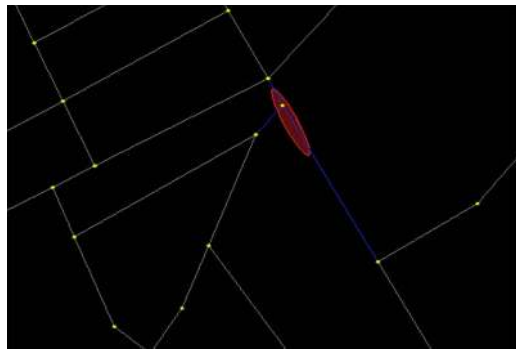


Fig. 3. (Red) Elliptical confidence region and (blue) candidate segments over the converted OpenStreetMap’s map of Alcalá de Henares.

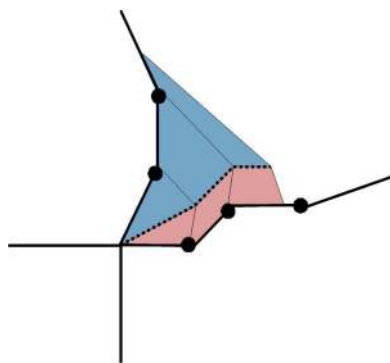


Fig. 4. Implemented curve-to-curve map-matching algorithm. The black points represent the nodes in the map, the thick lines are the road segments, and the dashed line is the estimated trajectory at the interjection exit. In shaded blue and red, the area from the trajectory curve to each one of the intersections exists. The road segment with the smaller area to the trajectory curve will be chosen.

If not, the distance from the motion trajectory to the segment is computed as follows.

- 1) Compute the starting point using the point-to-curve algorithm [41] for the last GPS position.
- 2) Compare the estimated run distance to the distance left on the starting point segment. If it is greater than the distance left, discard the starting point segment, and go to step 3. Otherwise, the starting point segment is the initial segment.
- 3) Compute the distance from the motion trajectory estimation to the candidate segments by computing the area under the motion estimation trajectory to a stretch of each one of the candidate segments. This stretch will have the same length as the motion trajectory estimation for all the candidate segments (see Fig. 4) [41].
- 4) Select the segment closer to the curve as the initial segment.

2) *Tracking of the Vehicle Position in the Map:* After setting the initial position of the vehicle in the map, subsequent motion estimations from the visual odometry are matched in the map, following a different approach. First, the vehicle velocity, heading, and position uncertainty are used to estimate if the vehicle is turning or driving through a junction. If so, the identification of the actual link is started. Otherwise, a simple tracking of the

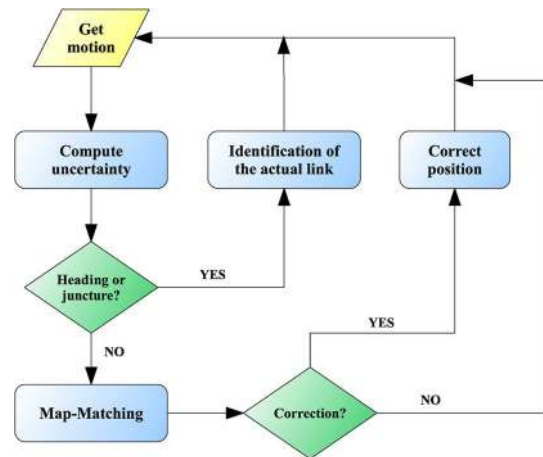


Fig. 5. Map-matching flow diagram.

vehicle position in the map is performed (see Fig. 5). The steps of this process are given here.

- 1) The heading of the vehicle and the current road segment orientation are continuously compared. If their difference is higher than 20° , a turning maneuver could have started so that the uncertainty in the position of the vehicle [see (7)] is incremented by 20%. Whenever the uncertainty region enters a road juncture, the speed and orientation of the vehicle are used to estimate if the car is turning. If so, the identification of the actual link is started; otherwise, it is continued.
- 2) Using the vehicle heading and velocity, check the position predicted by the visual odometry and the estimated position in the map after map matching. Whenever the map matching finds the vehicle position on a distinctive map feature (sharp turn, $> 50^\circ$ or roundabout), correct the vehicle position with the position in the map (correction); otherwise, just update with the new position.

This localization process is repeated until the GPS HDOP is less than 10; then, the GPS position will be used again.

V. EXPERIMENTAL SETUP

The system described was evaluated against data obtained during an experiment where a Citroën C4 was driven over 7 km through an urban setting in Alcalá de Henares, Spain. A stereo vision platform based on Basler sCA640 74-fm cameras was installed in the prototype vehicle. The stereo sequences were recorded using an external trigger signal for synchronization at 30 frame/s with a resolution of 640×480 pixels in grayscale. All sequences correspond to real traffic conditions in urban environments and speeds up to 50 km/h, with nonstationary vehicles and pedestrians. The vehicle has an on-board computer housing the image-processing system, an RTK-DGPS, and a low-cost GPS connected via USB (see Fig. 6). A specifically developed software captures the synchronized camera images and the RTK-GPS, GPS, and BusCAN information from the car. All this information is embedded into the 640×480 grayscale images in an overhead, along with the capture timestamp and the camera parameters (shutter, gain, exposure, etc.) for each image. The RTK-GPS receives



Fig. 6. (Top left) Stereo cameras. (Top right) RTK-GPS. (Bottom) Experimental vehicle.



Fig. 7. Flow diagram of the acquisition system.

differential corrections at 5 Hz from a base station through the Internet (see Fig. 7). Although corrections are sent at 5 Hz, the rover only computes the correction parameters at 1 Hz; therefore, we can consider that the RTK-GPS corrects at 1 Hz but delivers position information at 5 Hz. The low-cost GPS works at 1 Hz. The stereo sensor uses a baseline of approximately 300 mm and a focal length of 4.2 mm and is calibrated in an offline process. The trajectory ground truth is based on the position delivered by the RTK-DGPS, which has a measured maximum deviation in x - and y -axis of 5 and 5.6 mm, respectively, with standard deviations of 0.0036 and 0.0041 mm [7].

The digital maps can be downloaded at running time using OpenStreetMaps (OSM) [42] servers or can be downloaded offline and stored into a hard drive to be accessed on running time. Once the GPS signal is lost, the last reliable position is used to load an area of the OSM map surrounding that position. The map information is parsed and converted to Northing-Easting coordinates. All the conversions from and to WGS-84 latitude, longitude, and ellipsoid height to and from Universal Space Rectangular XYZ coordinates have been performed using the equations provided in [43]. This way, the motion information delivered by the visual odometry system in meters can be directly translated to the map, and our new position is estimated. Finally, our new position is converted back to longitude/latitude

and sent to a Java interface (Travelling Salesman [44]), which will represent our trajectory in the map.

Although much research has been performed in the area of visual odometry, there is no general way of performance assessment. In robotics and SLAM, a typical measure of the performance is the loop closure distance [9], [37], [45], which is computed as the distance to the starting point in a closed trajectory. The loop closure distance has the advantage that it does not require an external ground truth. It is strict with orientation errors and allows for correction when revisiting positions. On the other hand, scale errors might pass unnoticed, and small angular errors can accumulate, leading to very high loop closure distances. For this reason, the percentage of error in the distance traveled is also commonly used as an indication of the accuracy in the estimation of the translation [3], [18], but it requires a ground truth. When measuring the accuracy of the estimation of the rotation angles, the sensor used as ground truth is what usually determines the kind of performance measurement. Some works simply overlay the trajectory on a digital map and compare the estimated and manually annotated trajectory [25]. This method lacks quantitative measurement of the quality of the estimated trajectory. When using GPS devices, most of the works compare the trajectories in a bird's eye plot combined with the percentage of error in the distance traveled [3], [5], [18], [25]. With this method, small errors in the estimation of the angle can lead to big drifts, and the exact position and nature (impulsive, cumulative) of the error remains unnoticed. To overcome this, the works using inertial sensors (accelerometers, gyroscopes, and PHINS) usually plot the angular velocity against the estimated velocity [5]. This gives a good insight of the quality of the instantaneous estimation but masks the cumulative errors.

In this paper, we have tried to assess the quality of the system, giving both qualitative and quantitative measurements. For this reason, we will present overlaid trajectories of the estimated trajectory and the GPS-based trajectory, along with the percentage of error in the distance traveled, loop closure error, and root-mean-square error (RMSE) in positioning. To assess the correction capacity of the map-matching algorithm, the RMSE in positioning after the map matching has been used. The system limits has been tested, and examples of failure modes are also provided.

VI. RESULTS

A. Visual Odometry

The results of the visual odometry system alone are summarized in Table I, along with ground truth data from the GPS. The quality of the distance estimation is measured with the error in the distance traveled computed by the GPS and the visual odometry (columns 1 and 2), whereas the quality of the shape and angles of the trajectory are measured by the overlaid trajectories in Fig. 8 and the RMSE in positioning in Table I. The RMSE in positioning has been computed as the root-mean-square difference of the GPS position and the visual odometry position. The RMSE after map matching is computed using the corrected map-matching position, instead of the visual

TABLE I
GROUND TRUTH AND ESTIMATED LENGTHS USING VO ALONE

	GPS (m)	VO (m)	Distance Error (%)	RMSE (m)	RMSE After MM (m)	Video Time (mm:ss)	Process Time (mm:ss)
video 00	1098.1	1101.5	0.3	8.5	6.44	3:01	01:36
video 01	257.78	257.62	0.06	6.3	7.02	1:10	01:06
video 02	633.31	630.94	0.37	29.62	7.28	2:17	02:06
video 03	357.71	361.86	1.14	52.2	6.84	1:43	01:46
video 04	167.19	165.55	0.99	6.2	8.21	1:40	01:33
video 05	643.96	653.7	1.48	12.02	8.89	2:18	01:46
video 06	509.38	507.89	0.29	25.75	7.07	2:03	01:37
video 08	331.8	331.5	0.09	7.04	9.46	1:17	01:22
video 13	238.19	260	8.38	25.18	7.37	1:31	01:45
video 14	1147.8	1133.6	1.25	127.8	8.59	4:01	05:30
video 15	402.55	398.76	0.95	23.62	6.09	1:44	02:32
video 16	322.35	324.75	0.73	9.72	4.76	1:41	02:14
video 17	697.11	668.51	4.27	58.38	8.30	3:00	03:11
video 18	365.6	357.2	2.35	13.29	7.21	1:06	01:20
TOTAL	7172.83	7153.38	1.48			28:32	29:24

odometry position. The overall accuracy in the trajectory estimation is very high, specially taking into account that they were recorded under real traffic conditions with many nonstationary vehicles crossing the scene and strong glaring and shadows and tunnels. The mean error in the distance traveled is 1.48% for more than 7 km and almost 30 min of videos. To our knowledge, this is the best accuracy for such long and complex experiments under extreme illumination conditions and nonstationary objects. The complexity of the environment can be seen in the RMSE, where most of the errors can be explained by nonstationary objects occluding most of the image. In videos 3 and 13, three buses cross perpendicularly to the vehicle trajectory, occluding about 90% of the image and leading to badly wrong rotation estimations [see Fig. 9(b)]. In video 14, two cars slowly cross in front of the vehicle as handling a roundabout leading to a bad estimation in the exit angle from the roundabout. In videos 2, 6, and 17, the shadow casted by buildings make the images textureless, and very few and far features are available, increasing the error in the estimation [see Fig. 9(a)]. In videos 5 and 18, a tunnel is traversed, where very few features can be extracted and where the error in the estimation increases [see Fig. 9(c)].

In Fig. 8(a), the VO results, RTK-GPS, and GPS ground truth for a loop closure of 1.1 km are shown. In this case, the input video images are of poor quality as a result of glares on the windscreen and dazzling of the cameras. Several cars and one pedestrian crossed with the test vehicle while driving. Still, the system was able to correctly reconstruct the trajectory with an error in the loop closure of 0.2%. In Fig. 8(d), an example of error in the estimation caused by nonstationary objects (two cars and a bus) creates errors in the estimated exit angle of a roundabout, and the estimation drifts away from the real position. Still, the error in the estimated length (1.25%) proves that the errors are glitches. In the next section, this drift will be corrected using the map-matching algorithm. In Fig. 9, some frames of the failure modes are shown. Results show that the visual odometry is robust to illumination changes with good

results in underexposed and overexposed images. In these situations, the system is able to track very few features (5–10) and is more prone to errors. The small number of features available makes the motion estimations more noisy but is usually correct due to the restrictions introduced in the motion model and to the outlier rejection of RANSAC. On the other hand, the only defense against nonstatic points is the motion model and the assumption that most of the tracked points are static. Although this is enough for most of the situations, there are some cases, as shown in Fig. 9(b), where most of the points in the image are nonstatic and the only chance of avoiding a bad estimation is that the motion model of the system do not fit the motion of the nonstationary objects.

Results for very complex scenarios have been presented, including nonstationary cars, pedestrians, overexposed and underexposed images, and tunnels. However, the cumulative nature of the errors in visual odometry systems makes necessary a correction if global localization for longer runs wants to be achieved.

B. Visual Odometry and Map Matching

As we have shown in the previous section, the cumulative nature of errors in visual odometry and the numerous nonstatic objects of urban environments makes the trajectory estimations nonreliable for long runs. Here, we present the results of the proposed map-matching algorithm using the visual odometry motion trajectory estimation as input and corrected using the digital map features. In Table I, the quantitative results of the drift correction using the digital map features are presented in column *RMSE After MM*. The RMSE in positioning after the map matching is reduced when the estimation in the visual odometry step drifts. There is still a residual error after the map matching due to the *lock-on-road* feature, which forces the vehicle to move along a straight line between map nodes, but this error is bounded and does not increase in time.

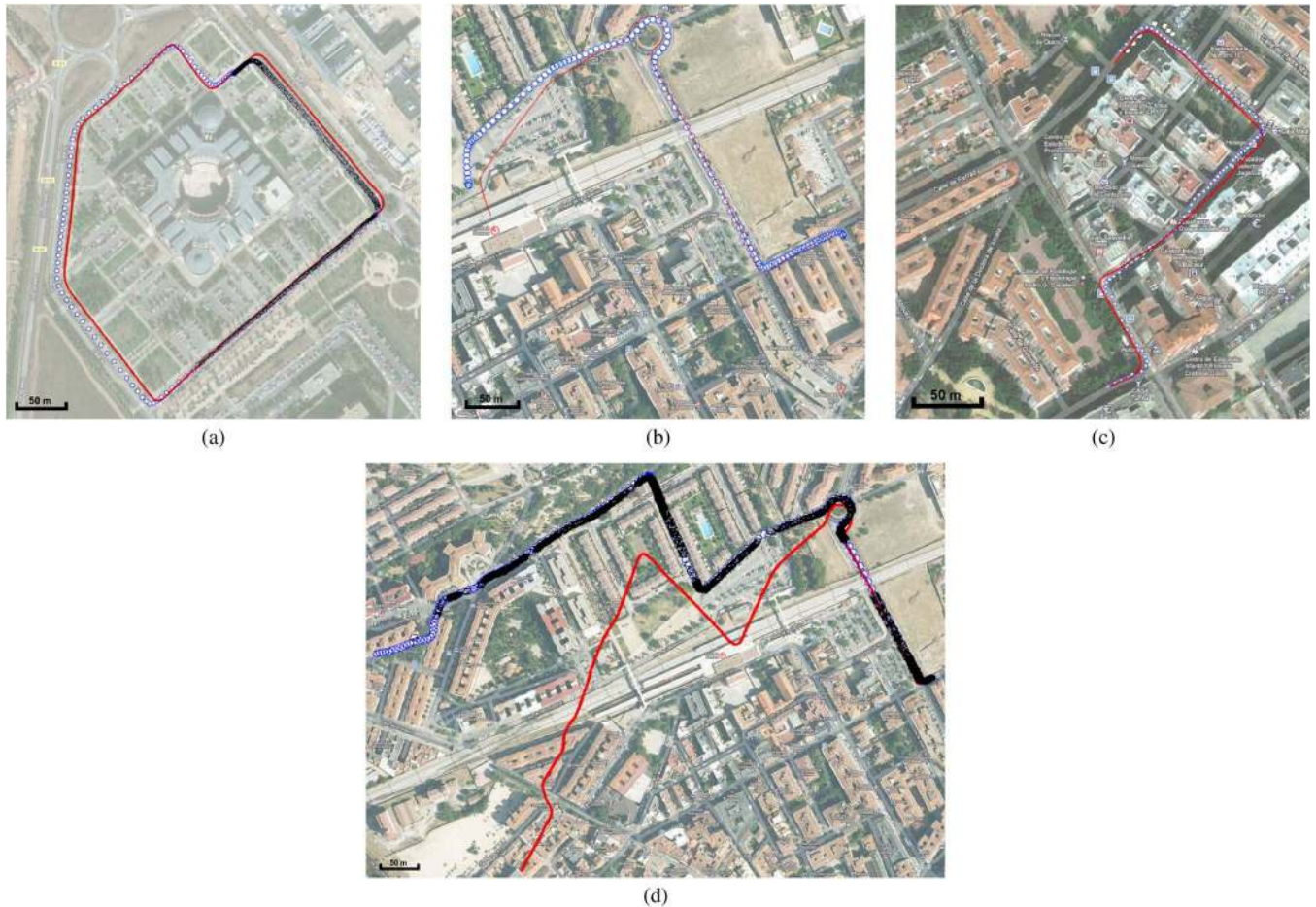


Fig. 8. Examples of 2-D estimated trajectories (blue circles, GPS; black diamonds, RTK-GPS; red line, VO). Upper row: Examples over 2 km with an accuracy of about 1%. Lower row: Example of failure mode due to illumination exceeding the dynamic range of the cameras and nonstationary objects. (a) Video 00: Closed loop, 1.1 km. (b) Video 5: Tunnel, 644 m. (c) Video 15: Urban, 402 m. (d) Video 14: Tunnel, 1.1 km.

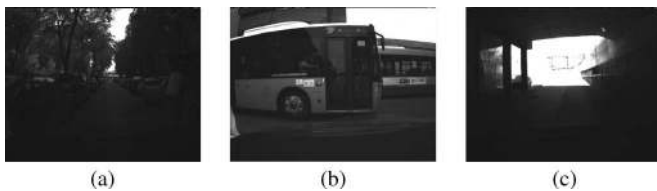


Fig. 9. Examples of frames where the system fails to get accurate estimations. (a) Example of underexposed image (Video 17, frame 879). (b) Example of nonstationary object failure (Video 13, frame 2377). (c) Example of tunnel (Video 5, frame 1331).

In Figs. 10 and 11, the global localization for two experiments of more than 1 km is shown as presented to the user. The map-matching algorithm output (latitude and longitude positions) is fed to the Java interface to OSM, i.e., Travelling Salesman [44], which performs the map rendering and trajectory representation. Information about the current position of the vehicle is presented in a navigatorlike interface in real time with information about the quality of the GPS signal, the number of satellites, speed, and altitude.

Comparing Fig. 8(d) and 10(b), it is shown that the global position of the vehicle is tracked with no mistakes, and the errors of the visual odometry are corrected by the

map-matching algorithm. The map-matching algorithm correctly estimates the exit for the roundabout where the visual odometry previously drifted away. Other cars and buses that are present on the video sequence that previously affected the visual odometry do not affect the global localization accuracy.

VII. CONCLUSION AND FUTURE WORK

A. Conclusions

We have described a new method for estimating the vehicle global position in a network of roads by means of visual odometry. To do so, a weighted nonlinear scheme, to represent the input data nature, has been explained and exhaustively tested. The resulting motion information has been used to get the global position of the vehicle in a digital map. Both systems have been used to compensate the GPS outages, performing global localization of the vehicle when the GPS signal is not available. The system has been implemented and tested under real traffic conditions. We have provided examples of the estimated vehicle trajectories and of the estimation of the global position in a map. Results show that the system is

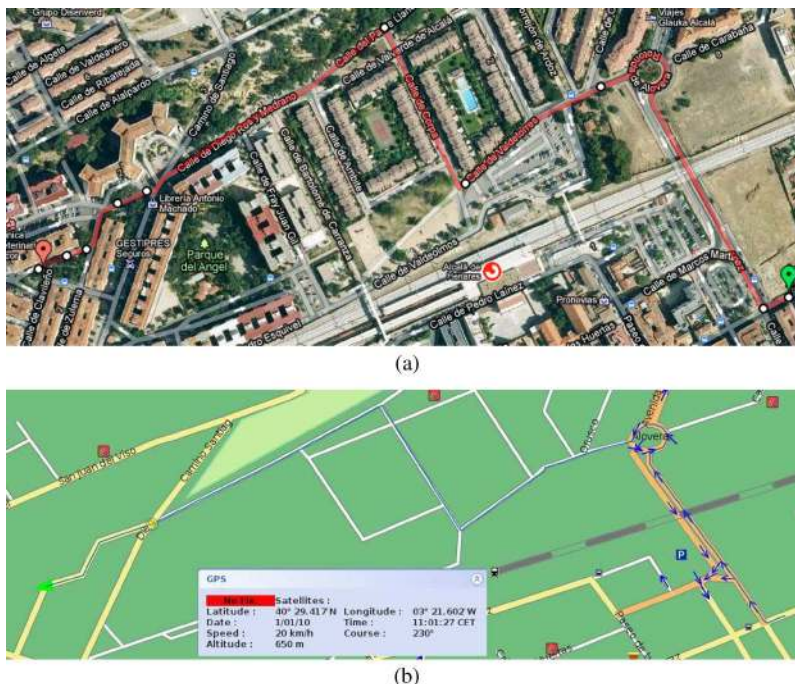


Fig. 10. Examples of reconstructed trajectories and information shown to the user for Video 14. (a) Trajectory of the vehicle for Video 14 overimposed on Google Maps. (b) Trajectory of the vehicle in blue and GPS information for Video 14, as shown to the user. The actual position of the vehicle is a green arrow.

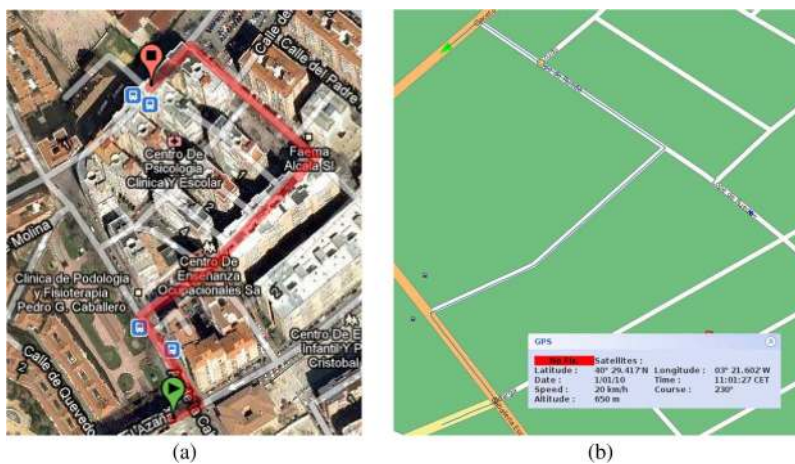


Fig. 11. Examples of reconstructed trajectories and information shown to the user for Video 15. (a) Trajectory of the vehicle for Video 15 overimposed on Google Maps. (b) Trajectory of the vehicle in blue and GPS information for Video 15, as shown to the user. The actual position of the vehicle is a green arrow.

capable of compensating the GPS outages and provides the global position to the user. The cumulative error from the visual odometry system is compensated using the topological information of the map. This indicates that longer outages can be corrected. The accuracy of the visual odometry estimation is 1.48% of the distance for more than 7 km and in very complex scenarios with illumination changes exceeding the dynamic range of the cameras and nonstationary objects. To our knowledge, this accuracy has never been reached before under these conditions. Moreover, a probabilistic map-matching algorithm constrained to the road, which uses map features to control the errors of the visual odometry by feeding back corrections from the map-matching process, has been proposed and tested. With this feature-based position correction, the RMSE error in positioning remains bounded, allowing the tracking of the

global position of the vehicle for very GPS outages on the order of tenths of kilometers.

B. Future Work

As part of our future work, we envision refining a method for discriminating stationary points from those which are moving in the scene [46]–[49]. Moving points can correspond to pedestrians or other vehicles. The ego-motion estimation will mainly rely on stationary points. The system can benefit from other vision-based applications currently under development and refinement in our laboratory, such as pedestrian detection [50], [51] and adaptive cruise control (based on vehicle detection) [52]–[54]. The output of these systems can guide the search for stationary points in the 3-D scene. In addition, longer runs of

tens or hundreds of kilometers have to be performed to estimate the correcting capability of the map fusion. Nighttime experiments have been performed, and we are currently analyzing the results and limits of the system. As part of future works, we intend to publish those results, along with adverse weather conditions.

REFERENCES

- [1] J. Ibanez-Guzman, S. Lefevre, A. Mokkadem, and S. Rodhaim, "Vehicle to vehicle communications applied to road intersection safety, field results," in *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst.*, Madeira Island, Portugal, 2010, pp. 192–197.
- [2] C. Wu, Y. Meng, L. Zhi-lin, C. Yong-qi, and J. Chao, "Tight integration of digital map and in-vehicle positioning unit for car navigation in urban areas," *J. Natural Sci.*, vol. 8, no. 2, pp. 551–556, 2003.
- [3] M. Agrawal and K. Konolige, "Real-time localization in outdoor environments using stereo vision and inexpensive GPS," in *Proc. 18th ICPR*, Aug. 2006, pp. 1063–1068.
- [4] A. Shashua, Y. Gdalyahu, and G. Hayun, "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," in *Proc. IEEE IV Symp.*, Parma, Italy, 2004, pp. 1–6.
- [5] S. Lovegrove, A. J. Davison, and J. Ibañez-Guzmán, "Accurate visual odometry from a rear parking camera," in *Proc. IEEE IV Symp.*, Jun. 2011, pp. 788–793.
- [6] N. Nourani-Vatani, J. Roberts, and M. Srinivasan, "Imu aided 3D visual odometry for car-like vehicles," in *Proc. IEEE Aust. Conf. Robot. Autom.*, Dec. 2008, pp. 1–8.
- [7] D. Fernandez Llorca, M. A. Sotelo, I. Parra, M. Ocaña, and L. M. Bergasa, "Error analysis in a stereo vision-based pedestrian detection sensor for collision avoidance applications," *Sensors*, vol. 10, no. 4, pp. 3741–3758, Apr. 2010.
- [8] A. Ishii, A. Sakai, M. Mitsuhashi, and Y. Kuroda, "Correcting angle of visual odometry system by fusing monocular and stereo methods in untextured dynamic environment," in *Proc. 10th Int. Symp. Commun. Inf. Technol.*, Oct. 2010, pp. 950–955.
- [9] D. Scaramuzza and R. Siegwart, "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles," *IEEE Trans. Robot.*, vol. 24, no. 5, pp. 1015–1026, Oct. 2008.
- [10] B. Matei and P. Meer, "Optimal rigid motion estimation and performance evaluation with bootstrap," in *Proc. CVPR*, 1999, pp. 339–345.
- [11] J. M. M. Montiel, J. Civera, and A. Davison, "Unified inverse depth parametrization for monocular SLAM," in *Proc. Robot.: Sci. Syst.*, Aug. 2006, pp. 932–945.
- [12] S. D. Blostein and T. S. Huang, "Error analysis in stereo determination of 3D point positions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 6, pp. 752–765, Nov. 1987.
- [13] F. Solina, "Errors in stereo due to quantization," Dept. Comput. Inf. Sci., Univ. Pennsylvania, Philadelphia, PA, Tech. Rep. MS-CIS-85-34, Sep. 1985.
- [14] C. Yuan and G. Medioni, "3D reconstruction of background and objects moving on ground plane viewed from a moving camera," in *Proc. IEEE CVPR*, 2006, pp. 2261–2268.
- [15] H. S. Park, T. Shiratori, I. Matthews, and Y. Sheikh, "3D reconstruction of a moving point from a series of 2D projections," in *Proc. ECCV: Part III*, 2010, pp. 158–171.
- [16] K. Konolige, M. Agrawal, and J. Solà, "Large scale visual odometry for rough terrain," in *Proc. ISRR*, 2007, pp. 201–202.
- [17] D. Nistér, O. Naroditsky, and J. Beren, "Visual odometry," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 2004, pp. 652–659.
- [18] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry for ground vehicle applications," *J. Field Robot.*, vol. 23, no. 1, pp. 3–20, Jan. 2006.
- [19] M. Sotelo, R. García, I. Parra, D. Llorca, M. Gavilán, S. Álvarez, and J. Naranjo, "Visual odometry for road vehicles—Fesibility analysis," *J. Zhejiang Univ.*, vol. 8, no. 12, pp. 2017–2020, Nov. 2007.
- [20] R. García, M. A. Sotelo, I. Parra, and D. F. Llorca, "3D visual odometry for road vehicles," *J. Intell. Robot. Syst.*, vol. 51, no. 1, pp. 113–134, 2008.
- [21] I. Parra, M. A. Sotelo, D. F. Llorca, and M. Ocaña, "Robust visual odometry for vehicle localization in urban environments," *ROBOTICA*, vol. 28, no. 3, pp. 441–452, May 2010.
- [22] H. Moravec, "Obstacle avoidance and navigation in the real world by a seeing robot rover," Ph.D. dissertation, Stanford Univ., Stanford, CA, Sep. 1980.
- [23] M. Maomone, Y. Cheng, and L. Matthies, "Two years of visual odometry on the mars exploration rovers: Field reports," *J. Field Robot.*, vol. 24, no. 3, pp. 169–186, Mar. 2007.
- [24] D. Nistér, "An efficient solution to the five-point relative pose problem," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2003, vol. 2, pp. 195–202.
- [25] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd, "Real time localization and 3D reconstruction," in *Proc. CVPR*, 2006, pp. 363–370.
- [26] D. Nistér, "A minimal solution to the generalised 3-point pose problem," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2004, vol. 1, pp. 560–567.
- [27] Z. Zhang and O. Faugueras, "Estimation of displacements from two 3D frames obtained from stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 12, pp. 1141–1156, Dec. 1992.
- [28] A. Hagnelius, "Visual odometry," M.S. thesis, Umea Univ., Umea, Sweden, 2005.
- [29] D. A. Forsyth and J. Ponce, "Computer Vision. A modern Approach," in *Pearson Education International*. Englewood Cliffs, NJ: Prentice-Hall, 2003.
- [30] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vis. Conf.*, 1988, pp. 147–151.
- [31] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Int. Joint Conf. Artif. Intell.*, Apr. 1981, pp. 674–679.
- [32] M. Agrawal, K. Konolige, and M. R. Blas, "Censure: Center surround extremas for realtime feature detection and matching," *Comput. Vis. ECCV*, vol. 5305, pp. 102–115, Oct. 2008.
- [33] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [34] P. Lothe, S. Bourgeois, F. Dekeyser, E. Royer, and M. Dhome, "Towards geographical referencing of monocular slam reconstruction using 3D city models: Application to real-time accurate vision-based localization," in *Proc. CVPR*, 2009, pp. 2882–2889.
- [35] A. Napier, G. Sibley, and P. Newman, "Real-time bounded-error pose estimation for road vehicles using vision," in *Proc. 13th Int. IEEE Conf. Intell. Transp. Syst.*, Madeira Island, Portugal, 2010, pp. 1141–1146.
- [36] A. C. Aitken, "On least squares and linear combinations of observations," *Proc. Roy. Soc. Edinburgh*, vol. 55, pp. 42–48, 1935.
- [37] G. P. Stein, O. Mano, and A. Shashua, "A robust method for computing vehicle ego-motion," in *Proc. IEEE IV Symp.*, 2000, pp. 362–368.
- [38] J. S. Greenfeld, "Matching GPS observations to locations on a digital map," in *Proc. 81st Annu. Meeting Transp. Res. Board*, Jan. 2002, p. 376.
- [39] M. A. Qudus, W. Y. Ochieng, and R. B. Noland, "Current map-matching algorithms for transport applications: State-of-the-art and future research directions," *Transp. Res. Part C: Emerging Technol.*, vol. 15, no. 5, pp. 312–328, Oct. 2007.
- [40] W. Ochieng, M. Qudus, and R. Noland, "Map-matching in complex urban road networks," *Brazilian J. Cartogr.*, vol. 55, no. 2, pp. 1–18, 2004.
- [41] D. Bernstein and A. Kornhauser, "An introduction to map matching for personal navigation assistants," New Jersey TIDE Center, Newark, NJ, Tech. Rep., 2002.
- [42] Steve Coast, Openstreetmaps, 2010. [Online]. Available: <http://wiki.openstreetmap.org>
- [43] S. H. Laurila and T. A. Stansell, *Electronic Surveying and Navigation*. New York: Wiley, 1976.
- [44] M. Wolschon, Travelling Salesman, 2010. [Online]. Available: http://wiki.openstreetmap.org/wiki/Travelling_salesman
- [45] L. Clemente, A. Davison, I. Reid, J. Neira, and J. Tardós, "Mapping large loops with a single hand-held camera," in *Proc. Robot.: Sci. Syst. Conf.*, Jun. 2007, pp. 1–8.
- [46] S. Avidan and A. Shashua, "Trajectory triangulation: 3D reconstruction of moving points from a monocular image sequence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 4, pp. 348–357, Apr. 2000. [Online]. Available: <http://dx.doi.org/10.1109/34.845377>.
- [47] M. Han and T. Kanade, "Reconstruction of a scene with multiple linearly moving objects," in *Proc. IEEE Conf. CVPR*, Jun. 2000, pp. 542–549.
- [48] A. P. Dani, Z. Kan, N. R. Fischer, and W. E. Dixon, "Structure and motion estimation of a moving object using a moving camera," in *Proc. Amer. Control Conf.*, 2010, pp. 6962–6967.
- [49] A. P. Dani, Z. Kan, N. R. Fischer, and W. E. Dixon, "Structure estimation of a moving object using a moving camera: an unknown input observer approach," in *Proc. 50th IEEE Conf. Decision Control*, 2011, pp. 5005–5010.

- [50] I. Parra, D. F. Llorca, M. Sotelo, L. M. Bergasa, P. Revenga, J. Nuevo, M. Ocaña, and M. A. García, "Combination of feature extraction methods for SVM pedestrian detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 292–307, Jun. 2007.
- [51] D. Fernandez Llorca, M. Sotelo, I. Parra, J. Naranjo, M. Gavilán, and S. Álvarez, "An experimental study on pitch compensation in pedestrian-portection systems for collision avoidance and mitigation," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 3, pp. 469–474, Sep. 2009.
- [52] F. J. Rodríguez, M. Mazo, and M. A. Sotelo, "Automation of an industrial fork lift truck, guided by artificial vision in open environments," *Auton. Robots*, vol. 5, no. 2, pp. 215–231, May 1998. [Online]. Available: <http://dx.doi.org/10.1023/A:1008826306614>.
- [53] Á. Sotelo, F. J. Rodríguez, and L. Magdalena, "Virtuous: Vision-based road transportation for unmanned operation on urban-like scenarios," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 2, pp. 69–83, Jun. 2004.
- [54] S. Álvarez, M. A. Sotelo, M. Ocaña, D. F. Llorca, I. Parra, and L. M. Bergasa, "Perception advances in outdoor vehicle detection for automatic cruise control," *ROBOTICA*, vol. 28, no. 5, pp. 765–779, Sep. 2009.



Sergio Álvarez Pardo received the B.Sc. degree in electrical and electronic engineering and the M.S. degree in advanced electronic systems, in 2006 and 2009, respectively, from the University of Alcalá (UAH), Alcalá de Henares, Spain, where he is currently working toward the Ph.D. degree with the Department of Computer Engineering, Polytechnic School.

His research interests include image processing, intelligent transportation systems, and robotics.



Miguel Ángel García-Garrido received the B.Sc. degree in industrial engineering, the M.Sc. degree in electronic engineering, and the Ph.D. degree in electrical engineering from the University of Alcalá (UAH), Alcalá de Henares, Spain, in 1998, 2001, and 2010, respectively.

He is currently a Lecturer with the Department of Computer Engineering, Polytechnic School, UAH. In 2009, he co-founded Vision Safety Technologies Ltd., which is a spin-off company established to commercialize computer vision systems. His

research interests are intelligent transportation systems, including advanced driver-assistance systems.



Ljubo Vlacic received the Grad. Diploma in engineering and the M. Phil. and Ph.D. degrees in electrical engineering (control), all from the University of Sarajevo in 1973, 1976, and 1986 respectively.

He is currently a Professor and the Director of the Intelligent Control Systems Laboratory, Institute of Integrated and Intelligent Systems, Griffith University, Brisbane, Australia.

Prof. Vlacic received the 2003 Queensland Professional Engineer of the Year from the Queensland Division of the Institution of Engineers Australia in recognition of his contributions to the advancement of control systems and their applications, the 2004 Sir Lionel Hooke Award from the Australian Council of the Institution of Engineering and Technology, the 2004 IEEE Achievement Medal (by the IEEE Knowledge Board), and a number of appreciation awards for notable services.



Miguel Ángel Sotelo (M'02) received the Dr.Eng. degree in electrical engineering from the Technical University of Madrid, Madrid, Spain, in 1996 and the Ph.D. degree in electrical engineering from the University of Alcalá (UAH), Alcalá de Henares, Spain, in 2001.

He is currently a Full Professor with the Department of Computer Engineering, Polytechnic School, UAH. He has been serving as an Auditor and Expert for the Fundación Instituto Tecnológico para la Seguridad del Automóvil Foundation, working on

R&D projects on automotive applications since September 2004. He is the author of more than 150 refereed publications in international journals, book chapters, and conference proceedings. His research interests include real-time computer vision and control systems for autonomous and assisted intelligent road vehicles.

Dr. Sotelo is a member of the IEEE Intelligent Transportation Systems (ITS) Society and the ITS-Spain Committee. He is currently an Associate Editor for the *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*. He received the Best Research Award in the domain of Automotive and Vehicle Applications in Spain in 2002 and 2009, the 3M Foundation Awards under the category of eSafety in 2003 and 2004, and the Best Young Researcher Award from the University of Alcalá in 2004.



Ignacio Parra Alonso received the M.S. degree in telecommunications engineering and the Ph.D. degree in electrical engineering from the University of Alcalá (UAH), Alcalá de Henares, Spain, in 2005 and 2010, respectively.

He is currently a member of the Research Staff with the Department of Computer Engineering, Polytechnic School, UAH. His research interests include intelligent transportation systems, intelligent vehicles, artificial vision, and operating systems.

Dr. Parra Alonso received the Master Thesis Award in eSafety the from the ADA Lectureship at the Technical University of Madrid, Madrid, Spain, in 2006 and the 3M Foundation Awards under the category of eSafety in 2009.



David Fernández Llorca (M'08) received the M.Sc. and Ph.D. degrees in telecommunications engineering from the University of Alcalá (UAH), Alcalá de Henares, Spain, in 2003 and 2008, respectively.

He is currently an Associate Professor with the Department of Computer Engineering, Polytechnic School, UAH. He is the author of more than 60 refereed publications in international journals, book chapters, and conference proceedings. His research interests are computer vision and intelligent transportation systems.

Dr. Fernández Llorca received the Best Ph.D. Award from UAH, the Best Research Award in the domain of Automotive and Vehicle applications in Spain in 2008, the 3M Foundation Awards under the category of eSafety in 2009, the Master Thesis Award in eSafety from the ADA Lectureship at the Technical University of Madrid, Madrid, Spain, in 2004, and the Best Telecommunication Engineering Student Award in 2004.



Miguel Gavilán received the M.Sc. and Ph.D. degrees in telecommunications engineering from the University of Alcalá (UAH), Alcalá de Henares, Spain, in 2007 and 2011, respectively.

He is currently a member of the Research Staff with the Department of Computer Engineering, Polytechnic School, UAH. His research interests include image processing and intelligent transportation systems.

Dr. Gavilán received the Master Thesis Award in eSafety from the ADA Lectureship at the Technical University of Madrid, Madrid, Spain, in 2007, the Master Thesis Award from the National Association of Telecommunication Engineers in 2008, and the 3M Foundation Awards under the category of eSafety in 2009.