This is a postprint version of the following published document:

Musleh, B.; Martín, N.; Armingol, J. M.; Escalera, A. (2014). Continuous pose estimation for stereo vision based on UV disparity applied to visual odometry in urban environments. *2014 IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, pp. 3983-3988.
DOI: 10.1109/ICRA.2014.6907437

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Continuous Pose Estimation for Stereo Vision based on UV Disparity Applied to Visual Odometry in Urban Environments

Basam Musleh, David Martín, José María Armingol and Arturo de la Escalera

Abstract— **This paper presents an autocalibration method to determine the pose of a stereo vision system based on knowing the geometry of the ground in front of the cameras. This pose changes considerably while the vehicle is driven, therefore it is good to know constantly the pose of the camera for several applications based on computer vision, such as advanced driver assistance systems, autonomous vehicles or robotics. These constant changes of the pose make interesting to be able to detect constantly the variations in its extrinsic parameters (height, pitch, roll). The validation of the autocalibration method is accomplished by a visual odometry implementation. A study of the improvement of the results of the visual odometry estimation taking into account the changes of the camera pose is presented, demonstrating the advantages of the autocalibration method.**

I. INTRODUCTION

Nowadays, advanced driver assistance systems (ADAS) and autonomous vehicles require complex applications in order to obtain accurate results. These applications use sensors such as lidar laser, monocular or stereo vision systems, among others. The ADAS algorithms relate the data provided by the sensors to the elements in traffic environments. In particular, the vision-based ADAS have to cope with the correspondence between the location of the objects in the world and its projection onto the image plane. This correspondence is determined by the intrinsic and extrinsic parameters of the camera. Intrinsic parameters are those related to the camera-optic set and can be determined by means of a calibration process. Extrinsic parameters express the pose of the camera: position (heigth h) and orientation (yaw ϕ, pitch θ and roll ρ) from the ground. These extrinsic parameters can change considerably while the vehicle is driven.

There are mainly three different ways to estimate the pose of the camera for traffic applications. The first one is by using a calibration pattern to determine the extrinsic parameters. For example, the calibration pattern can be placed on the ground [1][2], or painted in the hood of the vehicle [3]. Secondly, by way of making use of landmarks of the road [4][5], such as traffic lines [6], making easier the calibration process and allowing to be able to recalculate the extrinsic parameters at different times and places. This kind of landmarks cannot be detected constantly and the problems normally come up in urban environments, where the landmarks might be under bad conditions or occluded by other elements, for instance, parked cars. The last way is based on the geometry estimation of the ground in front of the vehicle [7] [8]. This geometry estimation makes possible to find out the pose of the camera, avoiding the necessity of a calibration pattern or landmarks.

This paper presents an autocalibration method to determine the extrinsic parameters of a stereo system based on knowing the geometry of the ground in front of the stereo rig. The disparity map [9] and the u-v disparity [10] [11] are used in order to distinguish between image points belonging to the ground (free map) and the ones which belong to the obstacles (obstacles map) [12]. Once the road has been detected, the goal of the autocalibration is to update the changes in the pose of the stereo rig regarding the road ahead of the vehicle. The implementation of the autocalibration allows to obtain measures of the height (h), pitch (θ) and roll (ρ) for each frame (Fig.1(a)). However, the yaw deviation (ϕ) is assumed constant in the implementation(Fig.1(b)). The autocalibration method has been applied to an algorithm of visual odometry [13][14]. Our visual odometry estimation [15] has been tested by using several trajectories of a public available dataset [16], which has ground truth information. The results of the visual odometry show the effect of each extrinsic parameter on the estimation of the vehicle movement.

The section II explains the autocalibration method, which is divided into the yaw calibration, and the calibration of the remaining extrinsic parameters. The application of the autocalibration method to a visual odometry algorithm is discussed in section III and its results are presented in section IV. Finally, the conclusions are put forward in section V.

II. AUTOCALIBRATION METHOD

The presented algorithm of autocalibration is divided into two stages. The deviation of the yaw (ϕ) between the stereo rig and the move direction of the vehicle is calculated in a unique calibration process, when the vehicle performs a straight line movement. The yaw deviation is assumed to be constant in time. The second stage consists in obtaining the heigth (h), the pitch (θ) and the roll (ρ) between the stereo rig and the road. This second stage is carried out frame by frame in order to detect the possible changes between the road and the stereo rig.

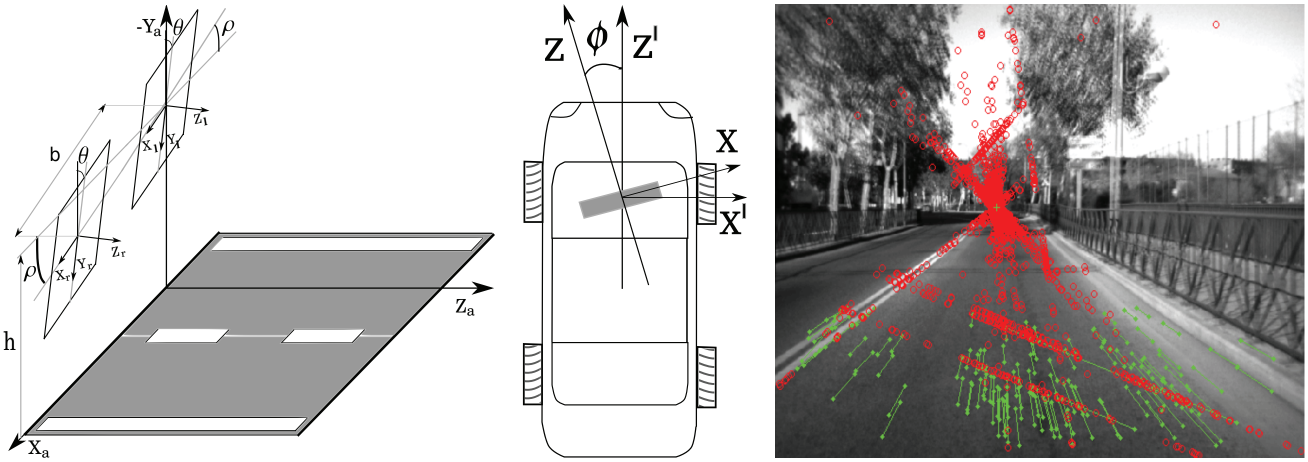
A. Yaw Calibration

The yaw calibration is based on detecting the vanishing point during two consecutive frames while the vehicle has a straight line motion. Ideally, when there is not a yaw deviation, the horizontal coordinate of the vanishing point (u_{vp}) should be equal to the horizontal coordinate of the

---

*This work was also supported by Spanish Government through the CICYT projects FEDORA (Grant TRA2010-20255-C03-01) and Driver Distraction Detector System (Grant TRA2011-29454-C03-02).

The authors are with the Intelligent Systems Lab at University Carlos III of Madrid, c/Butarque 15 28911 Leganes, Madrid, Spain.`{bmusleh, dmgomez, escalera, armingol}@ing.uc3m.es`



(a) (b) (c)

Fig. 1. (a) Schema of the configuration of the stereo rig from the ground. (b) Schema of the yaw deviation. (c) Results of an example of the yaw deviation calibration in an urban environment; Superposition of two consecutive frames where the consecutive positions of each feature point are connected by a straight line (green lines) whereas their intersections appear in red. Best viewed in colour.

optical center (u_0). Whenever a yaw deviation (ϕ) appears, it can be calculated using (1), where α is the focal length. The implemented method tracks feature points [17] between two consecutive frames in order to find the vanishing point. The coordinates of the nth feature point on the left image is (u_{n_1}, v_{n_1}) and (u_{n_2}, v_{n_2}) in the next frame. Thus, a straight line r_n (2) can be built for the feature point tracked. The vanishing point corresponds to the intersection of each straight line with the remaining lines.

$$\phi = \arctan\left(\frac{u_{vp} - u_0}{\alpha}\right) \tag{1}$$

$$v = \frac{v_{n_2} - v_{n_1}}{u_{n_2} - u_{n_1}} \cdot u + v_{n_2}(u_{n_2} + u_{n_1}) \tag{2}$$

The calibration process is performed in a non-controlled environment. Dynamic elements can appear, making difficult the determination of the vanishing point by means of this calibration process. The feature points belonging to the road are used to find the vanishing point, in order to deal with the dynamic elements. There are others sources of errors, such as the vehicle vibration or the incorrect matching of the feature points. An useful tool to solve these errors (outliers) is Random Sample Consensus (RANSAC) [18]. The resulting vanishing point for each pair of frames is chosen from every intersection points of the straight lines using RANSAC. This process is repeated along several frames and the final solution for the vanishing point is determined by using RANSAC again, among every partial solutions. An example of the processing of two consecutive frames is shown in Fig.1(c), where the position of each feature point in the two frames is connected by green lines. That is, the intersection of these lines determines the vanishing point.

B. Calibration of the Height, Pitch and Roll

Once the deviation of the yaw has been estimated (Fig.1(b)), the configuration of the stereo rig and the road is described in Fig.1(a). The stereo rig is composed of two cameras whose image planes are coplanar in order to make possible that the epipolar lines can be parallel. Following the nomenclature and explanation presented by Labayrade et al. in [7], it is possible to relate the homogeneous world coordinates of a point $P = (X, Y, Z, 1)$ to its projection in the two image planes $(u_i \cdot S, v \cdot S, S, 1)$ (3) and (4).

$$\begin{bmatrix} u_i S \\ vS \\ S \\ 1 \end{bmatrix} = M_{proj} \cdot M_{Translx} \cdot M_{Rotx} \cdot M_{Rotz} \cdot M_{Transly} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

$$\begin{aligned}
M_{project}(\alpha, u_0, v_0) &= \begin{bmatrix} \alpha & 0 & u_0 & 0 \\ 0 & \alpha & v_0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
M_{Translx}(\varepsilon_i b) &= \begin{bmatrix} 1 & 0 & 0 & -\varepsilon_i b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
M_{Rotx}(\theta) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
M_{Rotz}(\rho) &= \begin{bmatrix} \cos\rho & -\sin\rho & 0 & 0 \\ \sin\rho & \cos\rho & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
M_{Transly}(h) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & h \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned} \tag{4}$$

Where $i = r$ for the right camera and $i = l$ for the left one, being $\varepsilon_r = 1$ and $\varepsilon_l = 0$. b is the baseline between both cameras, (u_0, v_0) are the coordinates of the optical center and α is the focal length in pixels. The stereo rig is rotated an angle θ (pitch) around the axis X and an angle ρ (roll) around the axis Z, which correspond to the angles between stereo rig and ground (Fig. 1(a)). The cameras are elevated a height h from the ground.

The depth (Z) of any point of the world is a function of the disparity (Δ), which is the difference between the horizontal image coordinates of the point in both images. The expression of $u_r \cdot S$ for the right camera and $u_l \cdot S$ for the left one, is calculated by using (3). Then, the final expression for the disparity (Δ) for each point $P = (X,Y,Z,1)$ is (5).

$$\begin{aligned} \Delta &= \frac{u_l S - u_r S}{S} \\ \Delta &= \frac{\alpha \cdot b}{Z\cos\theta + (Y+h)\cos\rho\sin\theta + X\sin\rho\sin\theta} \end{aligned} \tag{5}$$

Below, the relationship between image coordinates for points which belong to the ground, is obtained. The ground or road in front of the stereo rig corresponds to every point whose $Y = 0$, as shown in Fig. 1(a). Thus, it is necessary to find out the expression of Y as a function of the coordinates of the image $\{u,v\}$, the extrinsic and intrinsic parameters of the stereo system $\{\alpha, b, u_0, v_0, h, \rho, \theta\}$, and the disparity ($\Delta$). The relationship between the world and left image ($\varepsilon_l = 0$) coordinates (3) can be formulated in inverse way (6), in order to obtain the expression for every point of the ground (7). The value of S is a function of the world coordinates (X,Y,Z), that is the main problem of this procedure and it is solved expressing S as $S = \alpha b/\Delta$ by using (5).

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = M^{-1}_{Tranly} \cdot M^{-1}_{Rotz} \cdot M^{-1}_{Rotx} \cdot M^{-1}_{Tranlx} \cdot M^{-1}_{proj} \cdot \begin{bmatrix} u\frac{\alpha b}{\Delta} \\ v\frac{\alpha b}{\Delta} \\ \frac{\alpha b}{\Delta} \\ 1 \end{bmatrix} \tag{6}$$

$$\begin{aligned} Y = -h + \frac{(v-v_0)b\cos\rho\cos\theta}{\Delta} - \frac{(u-u_0)b\sin\rho}{\Delta} + \\ + \frac{\alpha b\cos\rho\sin\theta}{\Delta} = 0 \Longrightarrow \end{aligned} \tag{7}$$

$$(v - v_0) = \frac{\tan\rho}{\cos\theta} \cdot (u - u_0) - \alpha\tan\theta + \left(\frac{h}{b\cos\rho\cos\theta}\right)\Delta \tag{8}$$

The equation (8) shows the relationship between the image coordinates (u,v) for the points belonging to the ground. This relationship is a straight line whose expression is $v = C \cdot u + d$ for the different values of disparity Δ. The roll in urban environments is normally low, so $\cos\rho \approx 1$ and $\sin\rho \approx 0$, simplifying the expression (8) and obtaining (9). This new expression (9) describes the relationship between the vertical coordinate image (v) and the disparity (Δ) as another straight line, that is achieved from the v-disparity. This straight line is known as road profile [10], being $v = C_r\Delta + v_{\Delta 0}$, where C_r is the slope and $v_{\Delta 0}$ is the value of (v) when the disparity is $\Delta = 0$. Once the road profile has been obtained, it is possible to calculate the pitch by means of (10) and the height h by using (11). The obtaining of the road profile when there are large obstacles -usually in urban environments- is a complex task. Thus, the road profile is obtained from a different v-disparity in the presented approach. This new v-disparity is built by using the free map instead of the disparity map [12], removing the obstacles.

$$v = \frac{h}{b\cos\theta} \cdot \Delta + v_0 - \alpha\tan\theta \tag{9}$$

$$\theta = \arctan\left(\frac{v_0 - v_{\Delta 0}}{\alpha}\right) \tag{10}$$

$$h = C_r \cdot b \cdot \cos(\theta) \tag{11}$$

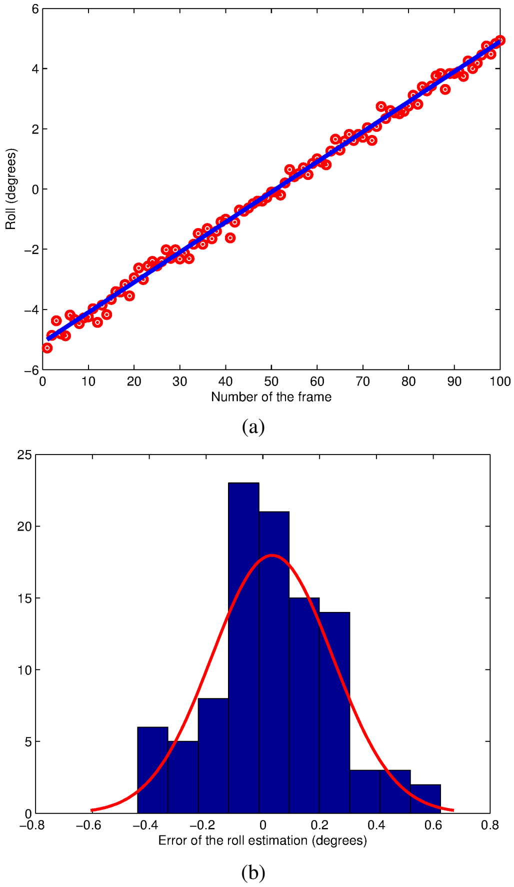
The next goal is to estimate the roll (ρ). The work presented by Labayrade et al. [7] performs several projections and regressions in order to estimate the roll. According to the authors, their implementation does not exceed 10 projections and regressions, but this implementation is computationally expensive. However, it is possible to determine the free map which corresponds to a disparity map, where only the ground appears in front of the vehicle, by means of the u-disparity [12]. Thus, the equation (8) can be applied to the free map to detect the straight line $v = Cu + d_\Delta$ for a fixed value of disparity Δ, which belongs to a close area of the vehicle. Then, the roll can be estimated knowing the slope C of the straight line in the free map by means of (12), avoiding carrying out several projections.

$$C = \frac{\tan\rho}{\cos\theta} \Rightarrow \rho = \arctan(C\cos\theta) \tag{12}$$

The evaluation of the roll estimation method is performed using a synthetic sequence of images. The ground appears in the synthetic images as a rolling plane which has been rotated in the range of -5 to 5 degrees. Fig.2 depicts the results of the evaluation, where the median and mean of the error are 0.0276 and 0.0331 degrees respectively, and the standard deviation is 0.213 degrees. The detail of the evaluation is shown in Fig.2(a), where the blue line corresponds to the gradual rotation performed over each frame, and the red dots are the corresponding results of the roll estimation method for each frame. Fig.2(b) shows the histogram and the normal distribution of the error in the roll estimation.

III. APPLICATION OF THE AUTOCALIBRATION TO VISUAL ODOMETRY

The presented approach for the autocalibration of the camera pose has been applied to our visual odometry method [15]. The estimation of the visual odometry is based on the fact that the vehicle performs a flat movement (2D) on the ground. This movement is estimated by tracking feature points [17] belonging to the road between consecutive frames of the left camera, and converting their coordinates from left image (u,v) into the world (X,Z). Once the image coordinates of the feature points have been obtained, the equation (6) is solved in order to express X (13) and Z (14) as a function of the coordinates of the image, the extrinsic and intrinsic parameters of the stereo system $\{\alpha, b, u_0, v_0, h, \rho, \theta\}$, and the disparity ($\Delta$). These world coordinates X (13) and



(a)

(b)

Fig. 2. Results of the roll estimation evaluation. (a) Comparison between the real rotation (blue line) and the estimated rotation (red dots) for each frame. (b) Histogram (blue columns) and the normal distribution (red lines) of the errors between the real rotation and the estimated one.

Z (14) do not have information about the calibration of the yaw deviation (ϕ). The solution to the yaw deviation is to calculate $X^{'}$ and $Z^{'}$ using (15) in order to estimate correctly the movement of the vehicle (Fig.1(b)).

$$X = \frac{b}{\Delta}(\cos\rho(u-u_0)+\cos\theta\sin\rho(v-v_0)+\alpha\sin\theta\sin\rho) \tag{13}$$

$$Z = \frac{b}{\Delta}(\alpha\cos\theta-\sin\theta(v-v_0)) \tag{14}$$

$$\begin{bmatrix} X^{'} \\ Z^{'} \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} \tag{15}$$

As we can see, both X (13) and Z (14) are a function of the disparity (Δ), whose value (5) is a function of the world coordinates (X,Y,Z). Thus, it is necessary to achieve a new expression of disparity (Δ) for the points belonging to the road. This new expression (16) can be obtained from (7) where the relation between the disparity and the image coordinate, the extrinsic and intrisic parameters is found.

$$\begin{aligned} \Delta = & \frac{b\cos(\rho)\cos(\theta)}{h}(v-v_0)-\frac{b\sin(\rho)}{h}(u-u_0)+ \\ & +\frac{\alpha b\cos(\rho)\sin(\theta)}{h} \end{aligned} \tag{16}$$

IV. EXPERIMENTAL RESULTS

The experimental results for the pose estimation of the camera and its effect on the visual odometry estimation are presented in this section. The evaluation of the presented approach has been performed by means of a publicly available dataset [16] [19], which allows to evaluate algorithms of visual odometry as the ground truth information is available. The presented results correspond to a sequence (2011_09_30_drive_0027) consisting of 1100 frames where a vehicle performs a closed loop in an urban environment. The trajectory has a lengh about 700 meters and a duration of 5 minutes.

A. *Results of the Camera Pose Estimation*

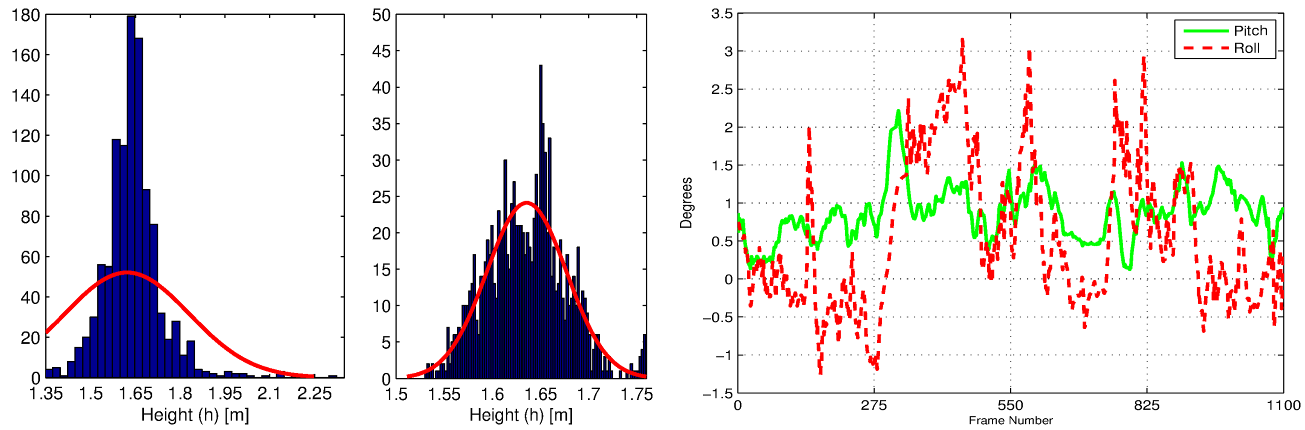
As we mentioned before, the camera pose estimation is divided into two stages: Firstly, the yaw deviation is achieved using the vanishing point (Section II-A). The yaw calibration has been performed by using different parts of sequences where the vehicle has a straight line motion obtaining a deviation value of -0.3 degrees. In spite of the fact that the yaw deviation seems to be low, it produces a great deformation on the resulting trajectory as it will be discussed in the next section (IV-B). Once the yaw calibration has been performed, and it is assumed constant along the trajectory, the next stage lies in obtaining the extrinsic remaining parameters, frame by frame, using solely the visual information.

The extrinsic parameter that best allows to evaluate the robustness of the autocalibration method is the height (h) where the stereo rig is located. The authors of the public dataset specify that the height has a value about 1.650 meters. The Fig.3(a) depicts the histogram and the normal distribution of the raw data for the height (h) estimated in the whole sequence (1100 frames). The results present a mean of 1.6404 meters and a median of 1.65 meters with a standard deviation of 0.0875 meters. The variations measured in the height can be due to several circumstances. Besides the possible outliers in the estimation of the extrinsic parameters, the height is a function of the load in the vehicle or the road due to potholes and speed bumps. By comparison, if the v-disparity is built by using the disparity map instead of the free map, the standard deviation increases to 0.1383 meters.

An Unscented Kalman Filter (UKF)[20] is used in order to reduce the number of outliers and smooth out undesirable fluctuations in the extrinsic parameters estimation (Fig.3(c)). The Fig.3(b) shows a new histogram and normal distribution of the filtered data for the height (h), reducing the standard deviation of the data to 0.0415 meters.

B. *Autocalibration Effect on the Visual Odometry*

The knowledge of both the yaw deviation (ϕ) and the continuous variations in each of the extrinsic remaining parameters (h,ρ,θ) is a useful source of information applied to the visual odometry estimation. This information allows to reduce in large measure the error in the visual odometry estimation. The Fig. 5 depicts the resulting trajectories of the visual odometry estimation, where it is possible to observe the effect of the autocalibration method in comparison with

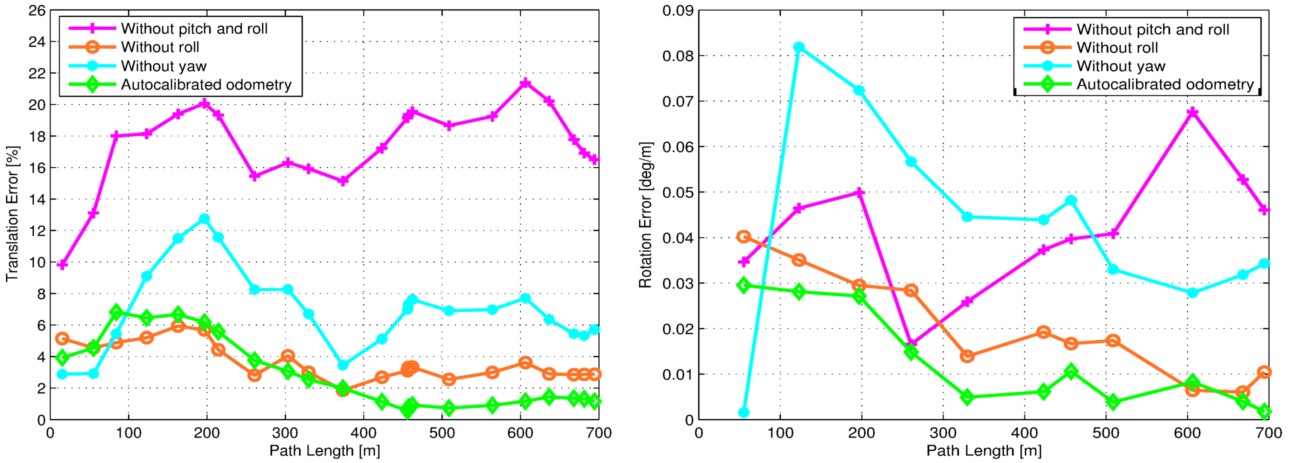


(a)

(b)

(c)

Fig. 3. Results of the estimation of the extrinsic parameters along the sequence: (a) and (b) Histogram and the normal distribution of the raw and filtered data respectively. (c) Evolution of the pitch (green line) and roll (dashed red line) all along the path.



(a)

(b)

Fig. 4. Comparison of the effect of the extrinsic parameters on the translation (a) and rotation (b) error evolution along the trajectory. Best viewed in colour.

the ground truth (red line). Besides, a quantitative analysis of the error both in translation and rotation is presented in Fig.4 and Table I, where the error values at the end of trajectory are also detailed.

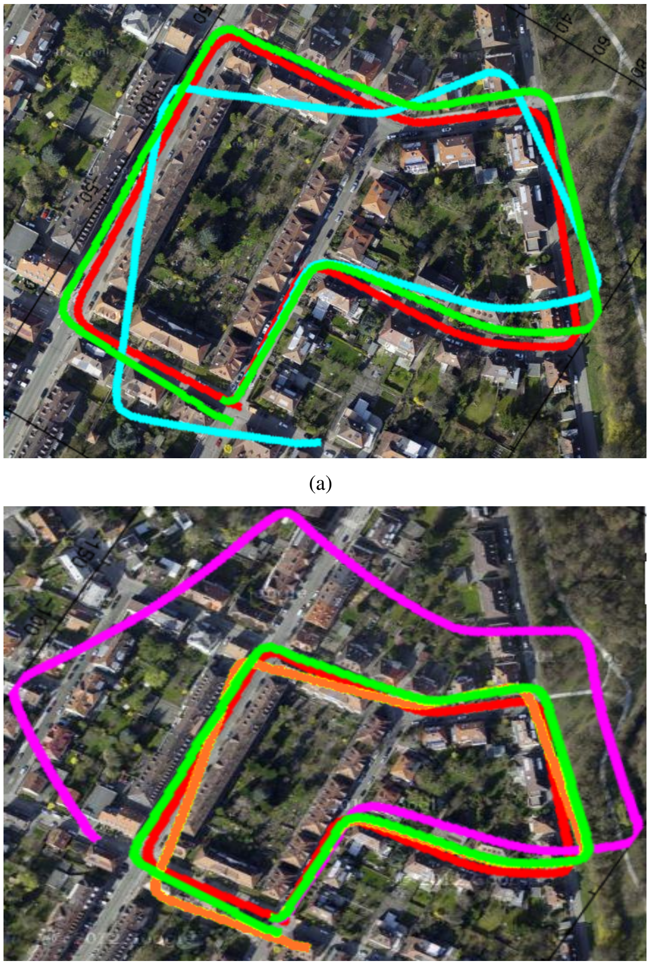
Firstly, the effect of the yaw calibration is shown in Fig. 5(a), where the resulting trajectory by using the information of every extrinsic parameters appears in green (Autocalibrated odometry), whereas the resulting trajectory appears in cyan if the yaw deviation is ignored ($\phi = 0$). A low yaw deviation produces an important rotation error along the trajectory (cyan line with dots in Fig.4) and, consequently, a large translation error at the end of the trajectory (5,72%).

In relation to the remaining angles of the extrinsic parameters (roll ρ and pitch θ), their effect on the trajectory is different. The autocalibration of the roll, frame by frame, improves to a lesser extent the estimation of the visual odometry than the autocalibration of the pitch. The Fig.5(b) puts forward a comparison of two resulting trajectories with the autocalibrated visual odometry (green line) and the ground truth (red line), where in the first one (line in orange), the roll has been fixed to ($\rho = 0$) and in the second one (line in magenta) the pitch has been also fixed to ($\theta = 0$). This different influence is also observed in the evolution of the error both in the rotation and the translation along the trajectory (Fig.4). Finally, the aggregation of the error because of the non-calibration of the roll and the pitch (in addition to the roll) produces a translation error of 2.87% and 16.51%, respectively. Whereas the autocalibrated estimation obtains a final translation error of 1.16% (Table I).

TABLE I

FINAL TRANSLATION AND ROTATION ERRORS OF THE TRAJECTORY

	Tran[m]	Rot[deg]	Tran[%]	Rot[deg/m]
no pitch & roll	114.70	31.96	16.51	0.0460
no roll	19.97	7.24	2.87	0.0343
no yaw	39.71	23.83	5.72	0.0104
autocalibrated	8.04	1.23	1.16	0.0018



(a)

(b)

Fig. 5. Resulting trajectories of the visual odometry estimation. (a) Comparison of the trajectory obtained without the information of the yaw deviation (cyan line) with the autocalibrated trajectory (green line) and the ground truth (red line). (b) Comparison of the effect of the roll (orange line) and the roll together with the pitch (magenta line) on the trajectories. Best viewed in colour.

V. CONCLUSIONS

An autocalibration method of the extrinsic parameters of a stereo rig, based on the road geometry in front of the vehicle, has been explained and applied to urban environments. The presented approach improves the method presented by Labayrade et al. [7], because of the road profile is obtained by using a v-disparity without obstacles, that reduces the number of outliers in the estimation. Besides, the roll can be achieved avoiding performing several projections with a precision about 0.03 degrees.

The application of the autocalibration method to a visual odometry algorithm has obtained a considerable reduction of the error along the trajectory followed by the vehicle, obtaining a final error of 1.16% in translation and a rotation error of 0.0018 degrees per meter. This reduction validates the performance and the utility of the autocalibration method in ADAS or autonomuos vehicles applications.

REFERENCES

[1] T. Marita, F. Oniga, S. Nedevschi, T. Graf, and R. Schmidt, “Camera calibration method for far range stereovision sensors used in vehicles,” in *Intelligent Vehicles Symposium, 2006 IEEE*. IEEE, 2006, pp. 356–363.

[2] S. Hold, C. Nunn, A. Kummert, and S. Muller-Schneiders, “Efficient and robust extrinsic camera calibration procedure for lane departure warning,” in *Intelligent Vehicles Symposium, 2009 IEEE*. IEEE, 2009, pp. 382–387.

[3] A. Broggi, M. Bertozzi, and A. Fascioli, “Self-calibration of a stereo vision system for automotive applications,” in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*, vol. 4. IEEE, 2001, pp. 3698–3703.

[4] S. Hold, S. Gormer, A. Kummert, M. Meuter, and S. Muller-Schneiders, “A novel approach for the online initial calibration of extrinsic parameters for a car-mounted camera,” in *Intelligent Transportation Systems, 2009. ITSC'09. 12th International IEEE Conference on*. IEEE, 2009, pp. 1–6.

[5] S. Li and Y. Hai, “Easy calibration of a blind-spot-free fisheye camera system using a scene of a parking space,” *Intelligent Transportation Systems, IEEE Transactions on*, vol. 12, no. 1, pp. 232–242, 2011.

[6] J. Collado, C. Hilario, A. de la Escalera, and J. Armingol, “Self-calibration of an on-board stereo-vision system for driver assistance systems,” in *Intelligent Vehicles Symposium, 2006 IEEE*. IEEE, 2006, pp. 156–162.

[7] R. Labayrade and D. Aubert, “A single framework for vehicle roll, pitch, yaw estimation and obstacles detection by stereovision,” in *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*. IEEE, 2003, pp. 31–36.

[8] Q. Wang, Q. Zhang, and F. Rovira-Mas, “Auto-calibration method to determine camera pose for stereovision-based off-road vehicle navigation,” *Environment control in biology*, vol. 48, no. 2, pp. 59–72, 2010.

[9] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” *International journal of computer vision*, vol. 47, no. 1, pp. 7–42, 2002.

[10] R. Labayrade, D. Aubert, and J. Tarel, “Real time obstacle detection in stereovision on non flat road geometry through v-disparity representation,” in *Intelligent Vehicle Symposium, 2002. IEEE*, vol. 2. IEEE, 2002, pp. 646–651.

[11] Z. Hu, F. Lamosa, and K. Uchimura, “A complete uv-disparity study for stereovision based 3d driving environment analysis,” in *3-D Digital Imaging and Modeling, 2005. 3DIM 2005. Fifth International Conference on*. IEEE, 2005, pp. 204–211.

[12] B. Musleh, A. de la Escalera, and J. M. Armingol, “U-v disparity analysis in urban environments,” in *Computer Aided Systems Theory EUROCAST 2011*, ser. LNCS. Springer, 2012, vol. 6928, pp. 426–432.

[13] D. Scaramuzza and F. Fraundorfer, “Visual odometry [tutorial],” *Robotics & Automation Magazine, IEEE*, vol. 18, no. 4, pp. 80–92, 2011.

[14] F. Fraundorfer and D. Scaramuzza, “Visual odometry: Part ii: Matching, robustness, optimization, and applications,” *Robotics & Automation Magazine, IEEE*, vol. 19, no. 2, pp. 78–90, 2012.

[15] B. Musleh, D. Martin, A. de la Escalera, and J. M. Armingol, “Visual ego motion estimation in urban environments based on uv disparity,” in *Intelligent Vehicles Symposium (IV), 2012 IEEE*. IEEE, 2012, pp. 444–449.

[16] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? the kitti vision benchmark suite,” in *Computer Vision and Pattern Recognition (CVPR)*, Providence, USA, June 2012.

[17] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.

[18] M. Fischler and R. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[19] KIT. (2013) The kitti vision benchmark suite. [Online]. Available: http://www.cvlibs.net/datasets/kitti/

[20] S. J. Julier and J. K. Uhlmann, “New extension of the kalman filter to nonlinear systems,” in *AeroSense'97*. International Society for Optics and Photonics, 1997, pp. 182–193.