This is a post-print version of the following published document:

Visual Ego Motion Estimation in Urban Environments based on U-V Disparity

Basam Musleh, David Martin, Arturo de la Escalera and Jose Maria Armingol

Abstract— The movement of the vehicle provides useful information for different applications, such as driver assistant systems or autonomous vehicles. This information can be known by means of a GPS, but there are some areas in urban environments where the signal is not available, as tunnels or streets with high buildings. A new method for 2D visual ego motion estimation in urban environments is presented in this paper. This method is based on a stereo-vision system where the feature road points are tracked frame to frame in order to estimate the movement of the vehicle, avoiding outliers from dynamic obstacles. The road profile is used to obtain the world coordinates of the feature points as a unique function of its left image coordinates. For these reasons it is only necessary to search feature points in the lower third of the left images. Moreover, the Kalman filter is used as a solution for filtering problem. That is, in some cases, it is necessary to filter raw data due to noise acquisition of time series. The results of the visual ego motion are compared with raw data from a GPS.

I. INTRODUCTION

Visual ego motion estimation or visual odometry are useful applications for developing Intelligent Transportation Systems (ITS), such as Driver Assistance Systems or autonomous vehicles. Several of these systems need to know the movement of the vehicle. This information is usually supplied by a GPS, but there are some areas where the signal is not available in urban environments. This is because the signal is affected by high buildings or tunnels. The method of visual ego motion estimation proposed in this paper allows to know the 2D displacements of the vehicle in urban environments with a great accuracy.

The visual ego motion is one of the most active field in computer vision. Different solutions have made use of monocular cameras [1], stereo systems [2] and omnidirectional cameras [3]. Monocular sensors have the problem of scalar factor, but its implementation is easier than stereo systems which presents a complex calibration step. Stereo systems achieve the most accurate results in long distances, because 3D information is avalaible, although they present great uncertainty in depth estimation [4]. Several methods have been presented, working in the disparity space in order to solve the uncertainty [5]. Finally, omnidirectional cameras allow to track the feature points along more frames than others sensors[6]. On the other hand, the distorsion makes difficult the feature matching.

The method presented is based on a stereo vision system [7]. The visual ego motion estimation or visual odometry is normally performed by means of detecting and tracking feature points between consecutive frames [8]. Our visual ego motion estimation uses a dense disparity map [9] to detect the road in front of the vehicle in order to use only the feature points that belong to the road, avoiding using feature points of obstacles which can be a source of outliers if the obstacles are moving. Another interesting result of using only feature points that belong to the road is that it is not necessary to search feature points in the whole image, as our algorithm only needs to process a lower third of the image.

The information of the road profile [10] is used to obtain the world coordinates of the feature road points. Besides the road profile, it is only necessary to know the position of the feature points on the left image in order to obtain the world coordinates, in contrast to most of the visual ego motion algorithms, which need to perform a matching between the images of the stereo pair in order to obtain the disparity for each feature point. Moreover, the feature points used are close to the vehicle reducing the uncertain in depth estimation.

Kalman filtering [11] has been applied to many situations in engineering, such as radio communication signals or applications to navigation. The filtering approach of the algorithm is applied in this work to raw data in order to smooth out undesirable fluctuations in visual ego motion.

An acquisition of raw data from a GPS is synchronized with the capture of images in order to compare the results of the visual ego motion with the GPS. The GPS is based on a Novatel receiver [12] that has been configured to work with Satellite-Based Augmentation Systems (SBAS) for sub-meter positioning. The solution has been calculated specifically with the European Geo-Stationary Navigation System (EG-NOS), which is a type of geo-stationary satellite system that improves the accuracy of the basic GPS signals. Accuracy is enhanced through the use of wide area corrections for GPS satellite orbits and ionospheric errors. EGNOS consists of a series of reference stations, master stations, ground uplink stations and geostationary satellites (GEOs).

II. OBSTACLES AND ROAD DETECTION

The dense disparity map and the u-v disparity [13] are developed in order to detect obstacles in front of the vehicle. The method for detecting obstacles obtains, as a result, two different dense disparity maps. The first one is the obstacle map (Fig.1(c)) and the second one, the free map (Fig.1(d)). The obstacle map is a disparity map where only

the obstacles appear, whereas the free map is the opposite to the obstacle map, where only the empty space ahead of the vehicle appears. This information can be used by a system of detection and localization of obstacles, as presented in [14].

A. Use of the Disparity Map and the u-v Disparity

Two cameras are neccesary to calculate the depth (Z) for a point $P = (X, Y, Z)$ in world coordinates by means of the stereo equation (1), where the projection of the point P over the image plane is (u_L, v_L) for the left camera and (u_R, v_r) for the right one.

$$Z = \frac{f \cdot B}{u_L - u_R} = \frac{f \cdot B}{d} \tag{1}$$

Where d is the disparity, f is the focal length and B is the baseline between both cameras. The disparity map is developed using the rectified images supplied by the stereo system, being the disparity (d) represented in the disparity map for every pixel of the image. The pixels corresponding to closest points have a bigger grey level and the farthest ones, an inferior grey level. Once the disparity map has been created, the u-v disparity can be obtained. The v-disparity expresses the histogram over the disparity values for every image row (v coordinate), whereas the u-disparity does the same, but for every column (u coordinate). It is important to note that a useful information can be obtained from the disparity map and its corresponding u-v disparity, such as the obstacles situated perpendicularly in front of the vehicle, which appear as horizontal lines in the u-disparity and as vertical lines in the v-disparity [15], in their corresponding values of disparity. As commented before, the obstacle map

(a) Left image   (b) Disparity map

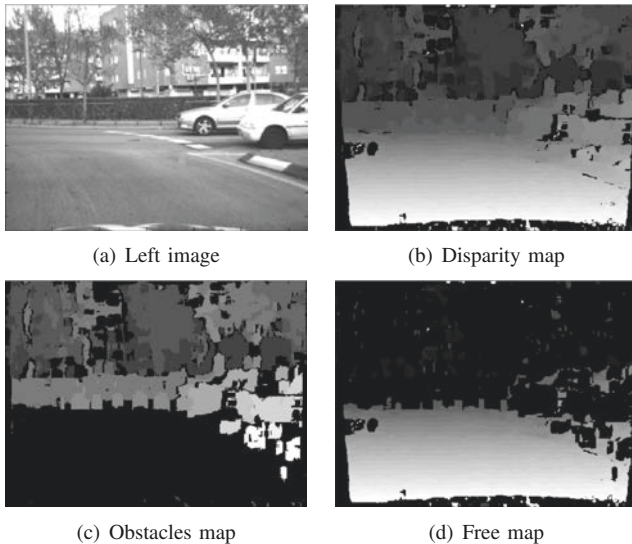(c) Obstacles map   (d) Free map

Fig. 1. Example of the Obstacles and Road Detection

is one of the results of the obstacle detection step. The development of the obstacle map is performed in two stages: firstly, the u-disparity is thresholded, using as a threshold the minimum height of an obstacle measured in pixels. This value must be small enough to detect every obstacle which blocks the movement of the vehicle and large enough

so as not to detect possible obstacles which do not avoid the movement of the vehicle, for example a speed bump. Once the u-disparity has been thresholded, every pixel of the disparity map which does not correspond to an obstacle is removed, getting the obstacles map. The second result of the obstacle detection is the free map. The development of the free map follows the same method that the generation of the obstacle map in the first stage, but in the second stage every pixel belonging to an obstacle is removed. Althought the free map usually corresponds to the road, the free map is theoretically the whole empty space ahead of the vehicle.

B. Estimation and Use of the Road Profile

In addition to the information of the obstacles, it is possible to obtain information about the road from the v-disparity. The road appears in the v-disparity as an oblique line called the road profile [10]. If a flat ground assuption is performed, this road profile can be expressed as a straight line (2), where v is the vertical coordinate of the image, m is the slope of the road profile and b is the theoretical value of the horizon of the stereo system. Furthermore, once the road profile has been estimated, it is possible to know the picht α between the stereo rig and the road [10] for each frame, by means of the (3) where C_v is the vertical coordinate of the optical center.

$$v = m \cdot d + b \tag{2}$$

$$\alpha = \arctan \frac{b - C_v}{f} \tag{3}$$

It is important to note that estimating the road profile v-disparity which has been generated using the disparity map, can be a difficult task in urban environments, so the detection of the oblique line corresponding to the road profile is difficult. An extended explanation of the solution to this problem is presented in [16]. For this reason, it is better to use the free map in order to generate the v-disparity because the obstacles are removed from the v-disparity, so it is easier to estimate the road profile.

The straightforward way to calculate the depth (Z) for a point is by means of the stereo equation (1), where the depth (Z) is a function of the disparity d. On the other hand, the road profile (2) shows the relationship between the vertical coordinate of the image v and the disparity d, then it is possible to obtain a new expression (4) merging the stereo equation (1) and the road profile (2) where now the depth (Z) is a function of the vertical coordinate of the image v, increasing significantly the resolution of the new equation [14]. This new expression to calculate the depth Z is only for points belonging to the road. Moreover, in order to calculate the depth to the vehicle instead of the depth to the camera it is necessary to use the information about the pitch α (3). Once the depth (Z) for a point belonging to the road has been calculated, it is possible to obtain its world coordinate X, which is a function of Z as (5) shows. Where C_u is the horizontal coordinate of the optical center.

$$Z = \frac{m \cdot f \cdot B}{v - b} \cdot \cos \alpha \tag{4}$$

$$X = \frac{Z \cdot (u - C_u)}{f} = \frac{m \cdot B \cdot (u - C_u)}{v \cdot b} \quad (5)$$

III. VISUAL EGO MOTION ESTIMATION

The visual ego motion estimation is based on tracking feature points between consecutive frames of the left camera. In opposition to the usual methods of visual ego motion estimation or visual odometry [7]; this method does not have to match up the feature points between both images of the stereo system to locate the points in world coordinates, because the method only uses the points belonging to the road. This points can be located on world coordinates by using (4) and (5) with the coordinates of points on the left image. An implementation of the Scale-Invariant Transform Feature (SIFT) detector and descriptor [17] developed for the MATLAB environment [18] is used in order to detect the feature points of the images.
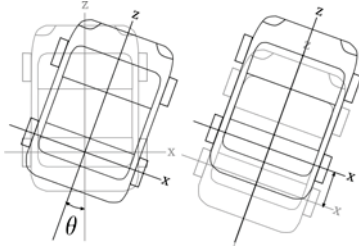
Fig. 2. Scheme of the movement of the vehicle. On the left, rotation stage. On the right, translation stage.

A. Vehicle Model

The chosen model for the kinematic motion of the vehicle is the Ackerman's steering model [19] which has been already used by other authors, such as in [20]. In order to simplify the visual ego motion estimation, it is necessary to make some assumptions: firstly, the movement of the vehicle between two consecutive frames can be divided into two stages, whose velocity is constant in each one (Fig.2): a rotation around the center of the motion of the rear axle and a forward translation after the rotation. The second assumption is that there is no slip in any direction.

B. Calculate of the Visual Ego Motion between Consecutive Frames

The vehicle motion estimation between two consecutive frames (t and $t + 1$) is perfomed in two steps. Firstly, the feature points have to be detected on the two left images of the stereo pair, in the instants t and $t + 1$. Then, a correspondence between the two sets of feature points (t and $t + 1$) is done in order to know the displacement of the feature points within the images. As commented before, the method makes only use of points of the road, which are situated at the bottom of the image. For this reason, it is possible to detect only the feature points at the bottom of the image, checking what feature points belong to the road using the free map. Besides the redution of the computing time, using only the closest points to the vehicle improves the flat ground assumption. Once the image coordinates of the

feature points have been obtained, it is possible to calculate the world coordinates $\{X, Z\}$ of these feature points in the instants t and $t + 1$ by means of (4) and (5).
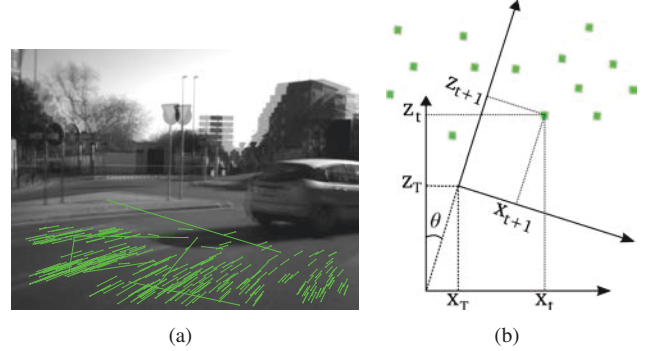
Fig. 3. (a) Example of feature points detecting and matching between consecutives frames. (b) Scheme of the movement of the cartesian coordinate system between consecutives frames.

$$\theta = \arctan\left(\frac{X_T}{Z_T}\right) \{ \quad \tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{X_T}{Z_T} \quad (6)$$

$$\begin{bmatrix} X_t \\ Z_t \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} X_{t+1} \\ Z_{t+1} \end{bmatrix} + \begin{bmatrix} X_T \\ Z_T \end{bmatrix} \quad (7)$$

Secondly, the estimation of the vehicle movement between two consecutives frames (t and $t + 1$) can be calculated by using the different locations in the instant t and $t + 1$ of the points detected in the previous step. As Fig.3(b) shows, the rotation angle θ of the vehicle can be calculated by means of the (6), where Z_T and X_T represent the translation after the rotation. Besides, it is possible to express Z_T and X_T as a function of θ and the locations of a road point, in the instant t and $t + 1$ by using (7). On the other hand, the angle θ is the only unknown variable of the expression in equation (8). It is possible to obtain firstly θ by solving the second order equation (9) and then Z_T and X_T by using (7).

$$\frac{X_T}{Z_T} = \frac{\sin\theta}{\cos\theta} = \frac{X_t - X_{t+1}\cos\theta - Z_{t+1}\sin\theta}{Z_t - Z_{t+1}\cos\theta + X_{t+1}\sin\theta} \quad (8)$$

$$(X_t^2 + Z_t^2)\sin^2\theta + (2 \cdot X_{t+1} \cdot Z_t)\sin\theta + (X_{t+1}^2 - X_t^2) = 0 \quad (9)$$

In this way, a set of solutions $\{\theta, Z_T, X_T\}$ for the visual ego motion estimation is obtained, where a solution $\{\theta, Z_T, X_T\}_k$ has been calculated for each pair of points $\{X_t, Z_t\}_k$ and $\{X_{t+1}, Z_{t+1}\}_k$. Different methods can be used in order to choose a unique solution $\{\theta, Z_T, X_T\}$ from the set of solutions, as a result of the visual ego motion between the two consecutive frames. Although the simplest method is the mean, the algorthim uses a solution from the set by means of the median because the median is more robust to possible outliers.

C. Model Using in Kalman filter

The process model is implemented by a linear time-varying (LTV) model in discrete time. That is, the process is described by a linear system. In this work, the system is a

vehicle driving along a road at constant velocity. This linear system is a process, which can be described by the following two equations:

$$x_{t+1} = A_t \Rightarrow x_t + w_t \qquad (10)$$

$$y_t = B \Rightarrow x_t + z_t \qquad (11)$$

Where, t is the time index, x is the state of the system and y is the measured output. The variable w is the process noise and z is the measurement noise. The matrix A is the state transition matrix and B is the measurement matrix, which are obtained to model a simple vehicle moving with constant velocity.

Then, the state vector x consists of the vehicle location $p = \begin{bmatrix} X, Z, \theta \end{bmatrix}$ and velocity v: $x_t = \begin{bmatrix} p_t \\ v_t \end{bmatrix}$, and the linear system equations are:

$$x_{t+1} = \begin{bmatrix} 1 & T_t \\ 0 & 1 \end{bmatrix} \Rightarrow x_t + w_t \qquad (12)$$

$$y_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \Rightarrow x_t + z_t \qquad (13)$$

The Kalman filter combines the measurements taken from the system (variables for the visual ego motion estimation $\}\theta, Z_T, X_T$), with the information provided by the motion model in order to obtain an optimal estimate of the system state. In this work, the measurement noise covariance matrix has been selected as the square of standard deviation of measurement (such as, 1^2 if we estimate Z_T), and the process noise covariance matrix uses process noise variance of 0.01.

IV. EXPERIMENTAL RESULTS

Several tests have been performed in urban environments in order to evaluate the goodness of the algorithm of the visual ego motion estimation. The results of a sequence, where the vehicle performes a closed loop in a urban environment, are presented in this section. The sequence has 1 minute and 38 seconds` duration and it is made up of 982 stereo images.

The Kalman Filter is applied to filter visual ego motion estimation and the Global Positioning System (GPS) is used to compare GPS raw data with visual ego motion estimation results.

A. Utility of the Kalman Filter

The visual ego motion estimation uses Kalman algorithm to solve filtering problem. The filtering approach of the algorithm is applied to raw data in order to smooth out undesirable fluctuations of visual ego motion variables $\}\theta, Z_T, X_T$.

The utility of the Kalman algorithm is displayed in Fig. 4, where graph represents an example of a curve trajectory of the vehicle. The example shows only one of the three variables corresponding to visual ego motion estimation, that is the angle θ. The example of the angle sequence starts at 20 second when the curve trajectory appears. The red circles of the graphs are visual ego motion angles and blue circles are obtained solution with Kalman algorithm.

The filter allows to smooth the fluctuations of the curve trajectory angle, eliminating outliers from visual ego motion estimations.
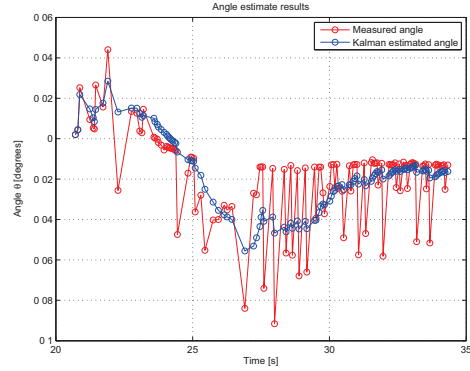

Fig. 4.   Example of filtering raw data

B. Results of the Visual Ego Motion Estimation

Two different methods are used in order to evaluate the degree of accuracy of the visual ego motion estimation. Firstly, it is possible to compare the difference between the initial and final location (position and orientation) of the vehicle. Due to the vehicle performes a closed loop, this two locations should be the same. Secondly, aerial imagery can be also used to overlay the resulting trajectory of the visual ego motion estimation.

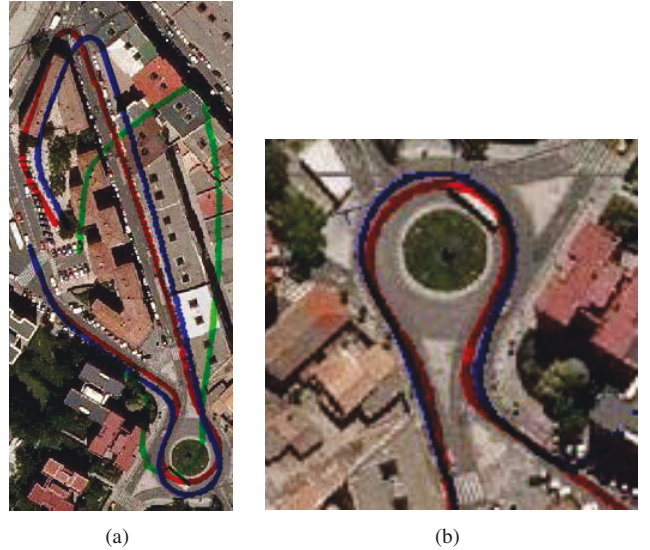
(a)   (b)

Fig. 5.   Comparison of the different trajectories of the visual ego motion estimation and GPS overlay in a aerial imagery. (a) With obstacles in blue, mean in green and final egomotion estimation (median) in red. (b) Detail for comparison of the visual egomotion estimation in red and the GPS raw data in blue.

Firstly, the two different methods (mean and median) to choose the final solution $\}\theta, Z_T, X_T$ from the set of possible ones $\}\theta, Z_T, X_T$ $_k$ are compared. Fig. 5(a) presents the resulting trajectory of the visual ego motion using both

4

methods: the trajectory calculated by the mean appears in green and the median in red. The trajectory is deformed due to use the mean, in comparison with the obtained one with the median. In fact, the difference between the position error is a 62.5% and the orietation error is a 173% so it can be concluded that it is better to use the median rather than the mean.

The effect of not using feature points from the obstacles has been studied. Fig. 5(a) shows the obtained trajectory using the feature points from the obstacles in blue. The fact of that using feature points from the obstacles ahead of the vehicle produce a 50% position error higher than the whole visual ego motion estimation, whereas the rotation error is similar. Regarding to the trajectory, the deviation is more prominent in the entire trajectory as Fig. 5(a) shows.

Finally, the raw data resulting from the visual ego motion estimation are filtered in order to smooth out fluctuations, Fig. 6 shows the results of filtering each variable $\}\theta, Z_T, X_T$. It is possible to observe the improvement of the smoothness in the three variables. The Fig. 6(a) displays the measured angle in red colour and Kalman estimated angle in blue colour. That is, Kalman algorithm estimates next angle using before angles of the time series. The results of the filter is appropriate in linear and curve trajectories, as can be observed in Fig.6(a), which first displays a linear trajectory followed by a curve to the left and so on. It is interesting to compare Fig.6(a) and Fig. 5(a) to observe the overall trajectory of the experiment, for example, when the vehicle is in the roundabout and how the Kalman filter smoothes continuously the curve trajectory to the left. The Fig.6(b) displays the X position and the behaviour of the filter is accurate obtaining again a smooth trajectory. It can be observed again the equivalence between Fig.6(b) and Fig. 5(a). The third graph is the Z position where the Kalman algorithm filters fluctuations of approximately 0.8 meters. The algorithm smoothes efficiently the Z variable.

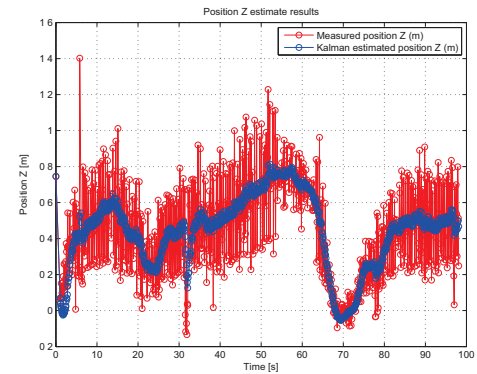### C. Comparision of visual ego motion Resultls and GPS

The GPS raw data of the followed trajectory by vehicle is displayed in Fig 7. Raw data is shown in Universal Transverse Mercator (UTM) geographic coordinate system to compare visual ego motion estimation data and GPS raw data in meters. The trajectory start, which is followed by vehicle, is marked with a red dot in Fig. 7, and each blue dot stands for a GPS point. The accuracy of overall trajectory is less than one meter when GPS + EGNOS is active, and if the receiver uses single point L1 solution the accuracy is 1.5 meters. Moreover, if the receiver is working in single point L1 solution and the number of GPS satellites is insufficient to calculate an optimum solution, then the accuracy of the solution can be increased more than 1,5 meters, resulting in a lateral displacement of the blue dots close to the real trajectory of the vehicle. Moreover, the trajectory shows some gaps, where the receiver can not compute the solution caused by the loss of GPS satellites. The loss of satellite signals is caused by obstructions from close buildings in the right and left side of the road. The final point of the trajectory
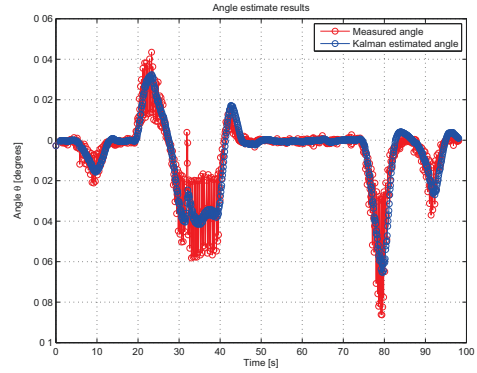
(a) Angle θ
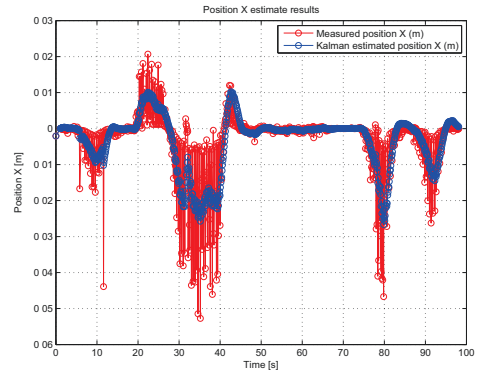
(b) Variable X

(c) Variable Z

Fig. 6. Results of the raw data filtering from visual ego motion estimation

is close to start point of the vehicle, and the difference is approximately one meter.

The comparison between visual ego motion estimation results and GPS raw data can be observed in Fig. 5(b). The visual ego motion trajectory is indicated by red colour and GPS raw data by blue colour. This comparison is the tool that allows establishing the performance of the visual ego motion estimation. Considering the accuracy of the GPS receiver, explained in former paragraph, the results establish that visual ego motion has better performance that GPS raw data. It is possible to observe wrong data at the bottom-right of the Fig.5(b) due to the receiver can not compute
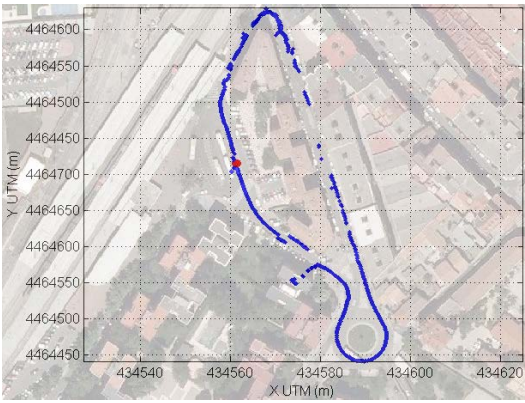
Fig. 7. Raw data of the GPS overlay over a aerial imagery

the solution caused by the loss of GPS satellites. The loss of satellite signals is caused by a close building in the right of the road. Moreover, the red trajectory matches exactly with the cars in the aerial image, as can be observed at the beginning of the roundabout where appears a car that is waiting for entering in the roundabout. A second example is in the middle of the roundabout, where a bus is performing a curve trajectory.

V. CONCLUSIONS AND FUTURE WORKS

The 2D visual ego motion estimation has been explained and applied in urban environments. The advantages of the smart algorithm have been shown comparing GPS raw data with visual ego motion results. The robust visual ego motion estimation ensures safe trajectory in case of GPS raw data loss caused by buildings, trees, tunnels, among other solid elements around vehicle. The GPS drawbacks have been solved with proposed algorithm. The visual ego motion estimation has reached the accuracy for curve and linear trajectories of the vehicle, avoiding outliers from dynamic obstacles. The results display a position error of 3.2% and a rotation error of 2.6% in a closed loop, which accomplishes the goal of the estimation algorithm. In addition, the use of the road profile and free map information allow only to search feature points in the lower third of the left images, and reduce the uncertain in depth estimation because these points are close to the vehicle. Kalman filter has been used as a great asset for smoothing out undesirable fluctuations of visual ego motion variables.

Regarding to the future work, a new method in order to autocalibrate the extrisic parameters will be implemented, allowing to change easily the position of the stereo rig. Another interesting improvement is the fusion of the visual ego motion estimation and the raw data of the GPS [21] in order to cope with GPS raw data loss and reduce cumulative error of the visual ego motion.

REFERENCES

[1] G. Stein, O. Mano, and A. Shashua, "A robust method for computing vehicle ego-motion," in *Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE*. IEEE, 2000, pp. 362–368.

[2] A. Howard, "Real-time stereo visual odometry for autonomous ground vehicles," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. Ieee, 2008, pp. 3946–3952.

[3] D. Scaramuzza, F. Fraundorfer, and R. Siegwart, "Real-time monocular visual odometry for on-road vehicles with 1-point ransac," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. Ieee, 2009, pp. 4293–4299.

[4] A. Hernández, J. Nieto, T. Vidal Calleja, E. Nebot, *et al.*, "Large scale visual odometry using stereo vision," 2011.

[5] D. Demirdjian and T. Darrell, "Motion estimation from disparity images," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1. IEEE, 2001, pp. 213–218.

[6] D. Scaramuzza and R. Siegwart, "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles," *Robotics, IEEE Transactions on*, vol. 24, no. 5, pp. 1015–1026, 2008.

[7] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1. IEEE, 2004, pp. I–652.

[8] I. Parra, M. Sotelo, D. Llorca, and M. Ocana, "Robust visual odometry for vehicle localization in urban environments," *Robotica*, vol. 28, no. 3, pp. 441–452, 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] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[12] http:www.novatel.com, NovAtel Inc., Calgary, Canada. (2012)

[13] N. Soquet, D. Aubert, M. Perrollaz, *et al.*, "Free space estimation for autonomous navigation," 2007.

[14] B. Musleh, A. Escalera, and J. Armingol, "Real-time pedestrian recognition in urban environments," *Advanced Microsystems for Automotive Applications 2011*, pp. 139–147, 2011.

[15] A. Broggi, C. Caraffi, R. Fedriga, and P. Grisleri, "Obstacle detection with stereo vision for off-road vehicle navigation," in *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*. IEEE, 2005, pp. 65–65.

[16] 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.

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

[18] A. Vedaldi, "An open implementation of the SIFT detector and descriptor," UCLA CSD, Tech. Rep. 070012, 2007.

[19] J. Borenstein, H. Everett, and L. Feng, "Where am i? sensors and methods for mobile robot positioning," *University of Michigan*, vol. 119, p. 120, 1996.

[20] N. Nourani-Vatani, J. Roberts, and M. Srinivasan, "Practical visual odometry for car-like vehicles," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 3551–3557.

[21] I. Parra, M. Angel Sotelo, D. Llorca, C. Fernández, A. Llamazares, N. Hernandez, and I. Garcia, "Visual odometry and map fusion for gps navigation assistance," in *Industrial Electronics (ISIE), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 832–837.