

POSITION ACCURACY ANALYSIS OF A ROBUST VISION-BASED NAVIGATION SYSTEM

S. Gaglione^{1*}, S. Del Pizzo¹, S. Troisi¹, A. Angrisano²

¹ Department of Science and Technology, University of Naples Parthenope, Napoli, Italy - (salvatore.gaglione; silvio.delpizzo; salvatore.troisi)@uniparthenope.it

² G. Fortunato University, Benevento, Italy - antonio.angrisano@unifortunato.eu

Commission II, WG II/1

KEY WORDS: Navigation; Space Resection; 3D map

ABSTRACT:

Using images to determine camera position and attitude is a consolidated method, very widespread for application like UAV navigation. In harsh environment, where GNSS could be degraded or denied, image-based positioning could represent a possible candidate for an integrated or alternative system. In this paper, such method is investigated using a system based on single camera and 3D maps. A robust estimation method is proposed in order to limit the effect of blunders or noisy measurements on position solution. The proposed approach is tested using images collected in an urban canyon, where GNSS positioning is very un-accurate. A previous photogrammetry survey has been performed to build the 3D model of tested area. The position accuracy analysis is performed and the effect of the robust method proposed is validated.

1. INTRODUCTION

Navigation solution in GNSS (Global Navigation Satellite system) denied or degraded environment has a growing demand. Satellite navigation has to be aided or replaced by other sensors in urban canyon or in-door scenarios. In such cases GNSS signals are blocked by artificial obstacles strongly degrading position accuracy or denying navigation (Angrisano et al., 2013). Other Radio Frequency (RF) signals are used for positioning where GNSS measurement are corrupted or insufficient. In in-door application Wireless Local Area Network (WLAN) based position methods are the most common. The preferred approach in signal degraded scenario is to use self-contained autonomous navigation as Inertial Navigation Systems (INS) integrated (Fisher, 2011a).

Another approach is the use of natural signal as magnetic field or light that provides measurement to many different non-GNSS methods. Visual sensors (e.g. cameras) play a key role in the last approach and are the main candidates to become an attractive substitute to GNSS thanks to the recent expansions in computational capability and in computer vision. Navigation system based on camera sensors are divided in two categories: vision-aided navigation system and vision-based navigation system. The first uses camera to provide measurement to aid a dynamic model of sensor motion (Indelman et al., 2012a; Indelman et al., 2012b). Vision-aided navigation produce an alternative and highly complementary system for constraining inertial drift. Vision-based navigation System, instead, localize body in-built with sensor processing only measurements provided by the camera-image. The analysis of image-scene provides significant landmarks enabling goal-seeking navigation. Among this category falls the structure for motion (SfM) and Bundle Block Adjustment (BBA) (Triggs et al., 2012; Indelman, 2012c; Engels, 2006). The common approach uses the images captured by camera to match against the expectation (prior knowledge of the navigation environment) in order to estimate camera position and attitude (Li et al.2011).

One challenge with navigation system based on natural signal, such as vision based, is the need for a reliable map, or "world model", which is required in order to make use of the natural measurements (Fisher et al., 2011b). This knowledge is fundamental when the navigation system is based on a single camera, like the case study. The use of 3-D maps of the navigational environment for camera positioning is possible considering photogrammetric method. Camera image is matched with 3-D map in real-time. The world coordinates of matched and identified points are transferred from 3-D map to current image; image coordinates of identified points are used as measurements and processed for camera positioning (Li et al.2011). The estimation method adopted in this study is the Space Resection Solution (SRS) (McGlone et al. 1980) based on Least Squared solution of measurement model. During the identification process, especially when performed automatically, outliers could occur and affect the position accuracy, reliability and, in worst case, cause solution failure.

In this study, a robust method for vision-based navigation system is proposed. In the space resection estimation process a robust approach is applied using a modified Huber M-estimator (Gaglione et al., 2017; Huber, 1981). In the Iteratively Reweighed Least-Squares (IRLS) a proper weighting matrix is applied to reduce the influence of leverage observables. The performance of proposed method is verified using images collected in an urban canyon scenario and results are showed.

The reminder of this paper is the following: in section 2 detail on the adopted vision-based navigation system and on the proposed estimation method are provided; section 3 is focused on experiments and results analysis; section 4 completes paper with conclusions.

2. VISION BASED NAVIGATION SYSTEM

2.1 Image Based Position Methods

Actually, localization services are in great demand. Of course, the traditional and advanced GNSS positioning techniques provide a reliable and accurate solution in outdoor environments, but their accuracy and reliability strongly decrease in urban canyon as well as in indoor positioning.

Promising image-based techniques allow obtaining the position and attitude of a camera. Several methodologies were developed in recent years, especially from the mobile robot community. These approaches can be divided into three main categories: Multi-view geometry; Stereo-view geometry; Single-view geometry.

The Multiview geometry is based on processing of image sequence, generally, the solution can be achieved in two different ways: in real-time or in post-processing. Strictly speaking the former approach continually estimates the camera position and attitude, for each image of the sequence. The camera localization is performed image by image. Generally, at the end of a single process such approach provides a pose estimation for each camera, therefore every time that a new image is added to the sequence the process estimates a new position for all images of the sequence (Wang et al., 2007). The most popular algorithm based on this type of approach is named SLAM (simultaneous localization and mapping), that allows to track the position of the camera during the navigation and to map unknown environments simultaneously (Bailey et al., 2006). The second approach is based on the classical photogrammetric procedures and algorithms, such as Bundle Adjustment, finding both the cameras positions and mapping in post processing (Troisi et al., 2017).

Stereo-view geometry is the oldest method employed to determine position of a robot in a known or unknown environment. Such approach would simulate the human vision system and it based on two cameras, constrained by a fixed distance. Generally, such approach is employed to estimate the relative position and orientation to an obstacle, in order to avoid it (Sabe et al., 2004).

Single view geometry is used to estimate the position and attitude of a camera in a specific reference system. Such methodology is generally based on the reference landmark and it estimates position and attitude of a camera image by image independently (Xun et al., 2011). Two popular algorithms are widely employed by the mobile robot community as well as by the computer vision and photogrammetric community: DLT (Direct Linear Transformation) and SRS (Space Resection Solution). Both allow to estimate the camera pose (position and attitude of a camera) with respect to a generic reference system, starting from the observation of several GCPs (Ground Control Points). Specifically, the DLT finds the solution setting up a linear equation model in the projective space, it needs at least six non-coplanar GCPs (Richard et al., 2003). On the other hand, the SRS provides the solution using three GCPs not aligned (McGlone, 1980). Other methodologies are applied to a single view geometry, such as a self-localization based on the relative position of the camera and straight lines recognized on both sides of the corridor (Hayashi et al., 2009).

In this work the single-view geometry approach has been employed to estimate the position of the camera. Specifically, such methodology is based on the Space Resection Algorithm and it introduces new approaches in order to increase the reliability and robustness of the methodology with outliers as well.

2.2 System Architecture

In this work, the experimentation approach is based on the comparison between the camera position solution, obtained using SRS, and classical topographic procedures. Basically, a photogrammetric survey of the scene was performed using several images, in according to the best practices recommended for this type of survey (Pierrot-Deseilligny et al., 2011) (Del Pizzo & Troisi, 2011), such as convergent images, a sufficient number of GCPs in order to limit the deformations and to obtain a georeferenced 3D model (Mikhail et al., 2001). Hence, the image acquired during the survey operations have been processed by a photogrammetric software using the classical procedure of Bundle Block Adjustment, which provides, as final results, a sparse 3D model of the scene and all images positions and orientations as well. Furthermore, image-matching operations were carried out in order to obtain a dense 3D point cloud (Figure 1).

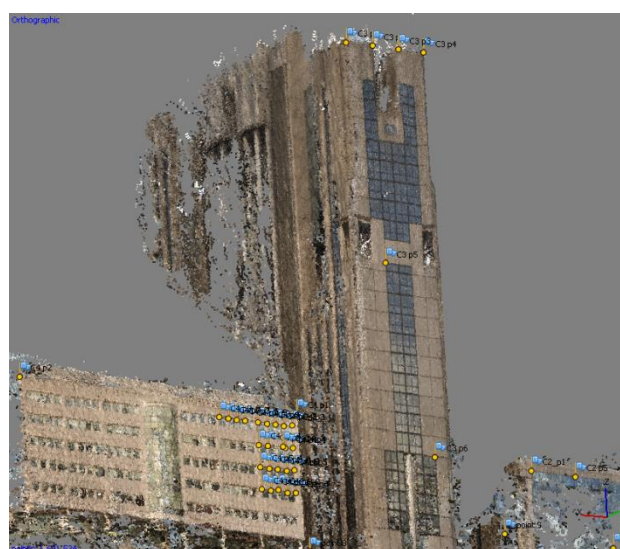


Figure 1. Raw dense 3D point cloud obtained from the post-processing of the photogrammetric survey. Blue flags are the GCPs used in SRS.

The obtained 3D model is scaled and georeferenced, it can provide further GCPs to use as input in developed SRS approach. Afterward, other images have been acquired in the same scenario to investigate SRS algorithm results.

Summarizing, the designed system provides two solutions for an image position: one obtained using the SRS approach, while the second provided by a classical topographic survey. Let assume the latter solution as reference, an analysis of SRS (using several approaches) error is carried out. Furthermore, synthetic outliers have been introduced in the GCPs coordinates in order to inspect the reliability of several SRS approaches.

The Figure 2 shows the designed system architecture employed for the inspection operations.

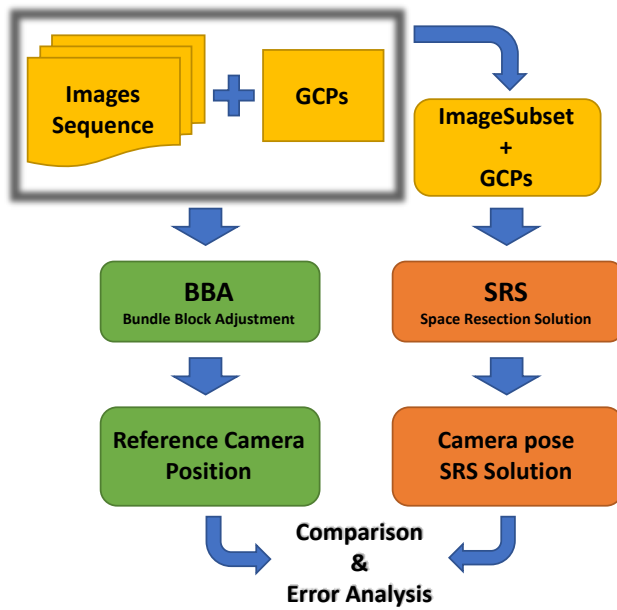


Figure 2. Designed system architecture to determine error analysis of SRS results.

2.3 Space Resection

The exterior orientation parameters of the camera, the sensor of the proposed vision-based navigation system, is computed by resection with respect to a 3D world reference frame. The resection is based on the linearization of collinearity equations (1), that describe the transformation of object coordinates (X, Y, Z) into corresponding image coordinates (x', y') .

$$\begin{aligned} x' &= x_0 + c \frac{r_{1,1}(X - X_0) + r_{1,2}(Y - Y_0) + r_{1,3}(Z - Z_0)}{r_{3,1}(X - X_0) + r_{3,2}(Y - Y_0) + r_{3,3}(Z - Z_0)} + \Delta x' \\ y' &= y_0 + c \frac{r_{2,1}(X - X_0) + r_{2,2}(Y - Y_0) + r_{2,3}(Z - Z_0)}{r_{3,1}(X - X_0) + r_{3,2}(Y - Y_0) + r_{3,3}(Z - Z_0)} + \Delta y' \end{aligned} \quad (1)$$

where $r_{i,j}$ are elements of the 3D rotation matrix R , performing a rotation from the image camera to the object reference system, while (X_0, Y_0, Z_0) are the coordinates of camera perspective point in object reference frame (Figure 3):

$$\begin{bmatrix} x' \\ y' \\ c \end{bmatrix} = R \begin{bmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{bmatrix} \quad (2)$$

The image coordinates are function of:

- the interior orientation parameters: $x_0, y_0, c, \Delta x', \Delta y'$; respectively: principal point coordinates, camera focal length and relative deviation due to distortion effects;
- the exterior orientation parameters $(X_0, Y_0, Z_0, \omega, \phi, \kappa)$, where last three elements are rotation angles: ω (tilt horizontal axis), ϕ (roll around azimuth axis), κ (roll around optical axes), as shown in Figure 3.

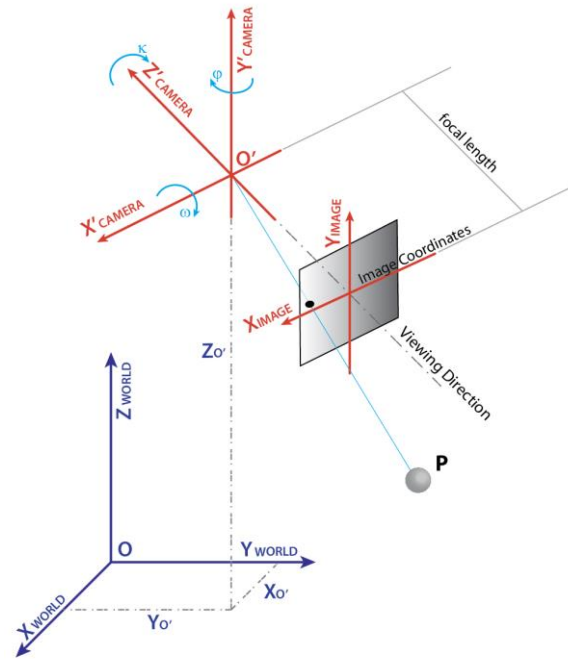


Figure 3. Two Reference Systems (RS): world-RS (O,XYZ) and the camera-RS (O',x'y'z').

The six exterior orientation parameters define the 6 degrees of freedom, i.e. 3 positions and 3 orientations of the vision sensor with respect to the world coordinate system.

The collinearity equation system (1) can be rewritten in the following system of correction equations:

$$\begin{aligned} x' &= x'_0 + vx' \\ y' &= y'_0 + vy' \end{aligned} \quad (3)$$

where (x'_0, y'_0) are the image coordinates computed using approximated exterior orientation parameters, while (vx', vy') are the adjustment parameters to obtain (x', y') , the correct image coordinates. The adjustment parameters can be obtained by linearization of the equation (1), using as initial point the approximated orientation parameters employed to compute (x'_0, y'_0) :

$$\begin{bmatrix} vx' \\ vy' \end{bmatrix} = \begin{bmatrix} \left(\frac{\delta x'}{\delta X_0} \right)_0 & \left(\frac{\delta x'}{\delta Y_0} \right)_0 & \left(\frac{\delta x'}{\delta Z_0} \right)_0 & \left(\frac{\delta x'}{\delta \omega} \right)_0 & \left(\frac{\delta x'}{\delta \phi} \right)_0 & \left(\frac{\delta x'}{\delta \kappa} \right)_0 \\ \left(\frac{\delta y'}{\delta X_0} \right)_0 & \left(\frac{\delta y'}{\delta Y_0} \right)_0 & \left(\frac{\delta y'}{\delta Z_0} \right)_0 & \left(\frac{\delta y'}{\delta \omega} \right)_0 & \left(\frac{\delta y'}{\delta \phi} \right)_0 & \left(\frac{\delta y'}{\delta \kappa} \right)_0 \end{bmatrix} \begin{bmatrix} dX_0 \\ dY_0 \\ dZ_0 \\ d\omega \\ d\phi \\ d\kappa \end{bmatrix} \quad (4)$$

The measurement model of the vision-based navigation system is obtained substituting (4) in (3):

$$\underline{z} = H \cdot \underline{\Delta x} + \underline{\varepsilon} \quad (5)$$

where the observation vector, the design matrix and unknown vector are:

$$\begin{aligned} \underline{z} &= \begin{bmatrix} x' - x'_0 \\ y' - y'_0 \end{bmatrix}, \\ H &= \begin{bmatrix} \left(\frac{\delta x'}{\delta X_0} \right)_0 & \left(\frac{\delta x'}{\delta Y_0} \right)_0 & \left(\frac{\delta x'}{\delta Z_0} \right)_0 & \left(\frac{\delta x'}{\delta \omega} \right)_0 & \left(\frac{\delta x'}{\delta \phi} \right)_0 & \left(\frac{\delta x'}{\delta k} \right)_0 \\ \left(\frac{\delta y'}{\delta X_0} \right)_0 & \left(\frac{\delta y'}{\delta Y_0} \right)_0 & \left(\frac{\delta y'}{\delta Z_0} \right)_0 & \left(\frac{\delta y'}{\delta \omega} \right)_0 & \left(\frac{\delta y'}{\delta \phi} \right)_0 & \left(\frac{\delta y'}{\delta k} \right)_0 \end{bmatrix}, \\ \Delta \mathbf{x} &= \begin{bmatrix} dX_0 & dY_0 & dZ_0 & d\omega & d\phi & dk \end{bmatrix}. \end{aligned}$$

Generally, the approximated position of the camera does not provide good results, therefore position and attitude are corrected using an iterative procedure converging to a solution after numerous iteration. At the end of iterative process, the method is able to estimate the camera attitude and position (Luhman et al., 2011).

In order to solve the measurement model (5), at least 3 points (not aligned) are necessary, indeed such observations provide 6 equations; the best estimation method for space resection is based on a Least Square (LS) solution of linearized collinearity equations with high level of accuracy in presence of redundant measurements.

LS estimation is extremely sensitive to outliers (Rousseeuw et al., 1987) and, in case of asymmetric error distribution, it can yield to an inaccurate final estimation.

2.4 Robust proposed method

Especially when the image coordinates are provided by automated methods, large errors among measurements could frequently occur, so a robust estimation approach has to be applied. The proposed robust method is a Huber M-estimator, enhanced with redundancy matrix information, in order to avoid the potential harmfulness of leverage observations (Leick, 2004). Fundamentals of the robust estimation theory were established by Box (1953) and later developed by other researchers (Hampel et al., 1986). Robust estimation methods are inherently resistant to gross errors, and some classes of them are based on the LS residuals:

$$\underline{r} = \underline{z} - H \cdot \underline{\hat{x}} \quad (6)$$

where $\underline{\hat{x}}$ is the LS estimation.

To reduce or remove the influence of gross errors, some class of robust estimators (e.g. M-estimators) adjust the measurement weights, using them in a Weighted Least Squares (WLS) approach to compute the solution (Huber, 1981).

Several robust estimators exist and belong to three main classes: M-estimators, L-estimators and R-estimators (Hampel et al., 1986). In this paper, the first class was chosen, considering the results obtained by previous analysis from application in GNSS context (Gaglione et al., 2017).

M-estimators have been introduced by Huber (1964) as a generalization of maximum likelihood estimators. It is related to least squares procedures and minimizes the objective function ρ as follows:

$$\min \sum_{i=1}^m \rho(r_i) = \min \sum_{i=1}^m \rho(z_i - H_i \cdot \underline{\hat{x}}) \quad (7)$$

where m is the number of the observations, r_i and z_i are respectively the i -th element of residuals and the uncorrelated measurements vectors, H_i is the i -th row of the design matrix H , and $\underline{\hat{x}}$ is the state vector.

The method applied in this study is an IRLS (Wieser et al., 2002) where weights depend on the residuals. According to this approach, starting from an initial WLS estimate $\underline{\hat{x}}$, residuals are computed and weights are adjusted, then a new WLS estimate is solved, and the process is repeated until the convergence is achieved.

In order to adjust weights at each iteration, Huber's M-estimator uses the following weight function:

$$(w_i)_{j+1} = \begin{cases} 1, & \frac{|(r_i)_j|}{\hat{\sigma}_0} \leq k \\ \frac{k}{|(r_i)_j/\hat{\sigma}_0|}, & \frac{|(r_i)_j|}{\hat{\sigma}_0} > k \end{cases} \quad (8)$$

where k is the tuning constant (set to 1.345), $\hat{\sigma}_0$ is the standard deviation of the residuals (Knight et al., 2009) and j indicates the iteration.

The Huber method provides the solution by changing the weights assigned to the observations at each iteration, hence w_i is a key parameter in the estimation. In this work, the a priori weight matrix W has been computed by using the redundancy matrix L :

$$L = I - H(H^T H)^{-1} H^T \quad (9)$$

The trace of the matrix L is the total redundancy (or the degree of freedom) of the measurement model, that is $(m - n)$, where m is the number of observations and n the number of unknowns; the i -th diagonal element l_i of L is called "redundancy number" and is the contribution of the i -th measurement to the total redundancy (Schaffrin, 1997).

The redundancy number assumes values between 0 and 1. Small values of l_i (near 0) correspond to measurements providing little contribution to total redundancy and so hardly controlled; on the other hand, approximately equal values of l_i are desirable, being every measurement controllable.

Measurements with small l_i values are leverage observations and have high potential to influence the solution; if a blunder or a large bias is present on a leverage observation, harmful effects can be evident on the positioning.

Furthermore, a modified approach of Huber method is adopted by the authors in order to take into account weights from the previous iteration (Yang, 1994). In detail, at each iteration, the weight matrix has been updated: the i -th diagonal element w_i of the new weight matrix W is determined as follows:

$$(w_i)_{j+1} = \begin{cases} (w_i)_j, & \frac{|(r_i)_j|}{\hat{\sigma}_0} \leq k \\ (w_i)_j \frac{k}{|(r_i)_j/\hat{\sigma}_0|}, & \frac{|(r_i)_j|}{\hat{\sigma}_0} > k \end{cases} \quad (10)$$

The current weight matrix is expressed as follows:

$$W = \text{diag}([w_1 \quad \dots \quad w_m]) \quad (11)$$

In the proposed approach, as initial weight in Huber iterative method, the diagonal element of the Redundancy matrix (9) is considered.

3. TEST AND RESULTS

3.1 Image Set-Up

In order to validate the robustness of the proposed approach on space resection two different images are considered (Figure 4). Pictures were collected in two different day-time (A-image during afternoon, B-image during the morning), and focus almost same buildings. The relative distance between camera and buildings is different in both images, respectively about 32 and 12 meters respectively. Different distances were chosen in order to evaluate the influence of the distance camera – object on position accuracy.



Figure 4. Two Processed Images

Both images were shot by a Nikon 800E camera with a 20mm lens. The lens was calibrated with a standard photogrammetric procedure in order to obtain the camera interior orientation parameters (Fraser, 1997; Faugeras et al. 1992, Remondino & Fraser, 2006). The image coordinates of the same 8 points were extracted from both images (Figure 5) (0).

Point ID	A-Image Coordinate		B-Image Coordinate	
	x (px)	y (px)	x (px)	y (px)
1	2558,4585	215,8849	3841,9673	1712,9646
2	2671,5542	225,9325	3996,6819	1870,3370
3	2780,8618	236,7529	4143,3608	2019,9237
4	2887,0322	245,6562	4287,5835	2168,2258
9	2701,5825	1407,0952	2907,3691	2922,7231
10	2685,2332	1482,6410	2748,7297	3001,4729
11	2736,7708	1628,0031	2641,7239	3336,3450
12	2790,8430	1787,9521	2505,8384	3760,3491

Table 1. Markers' Image Coordinates

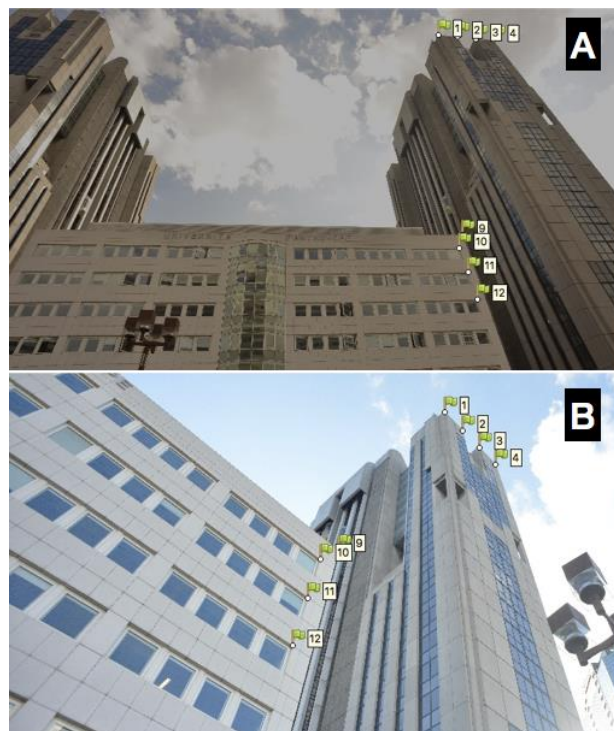


Figure 5. Points' Position on images

World coordinate of markers were obtained from the georeferenced 3D model (Table 2).

Marker ID	UTM Coordinate		
	Easting (m)	Northing (m)	Altitude (m)
1	439655,3769	4523112,669	142,488627
2	439650,6441	4523112,116	142,457771
3	439645,9769	4523111,576	142,368884
4	439641,3319	4523110,978	142,427165
9	439668,8666	4523114,812	85,107758
10	439669,84	4523114,942	83,190671
11	439669,8644	4523114,951	79,829578
12	439669,8754	4523114,976	76,471020

Table 2. Points' World Coordinates

3.2 Position accuracy analysis

The robust approach adopted in space resection is detailed in section 2 and it is based on the combined use of a robust estimation method, Huber M-estimator, with a proper initial weight matrix, the redundancy matrix. Also the single contribution of these single method is studied. So the compared configurations, implementing space resection, are:

- The one using as estimation method a WLS with diagonal element of redundancy matrix as weights, indicated as RLS;
- The one using as estimation method a Huber IRLS, indicated as HIRLS;

- The one based on the combined use, that represent the proposed approach, indicated as WHIRLS

The previously mentioned configurations are compared with the baseline one, where LS is the adopted estimation method.

The figure of merits adopted for the comparison are the horizontal and vertical components of the position error that is computed using a very accurate camera position. This reference is obtained by a topographic survey in case of B-image while for A-image it was used the bundled adjusted camera position obtained during the 3D model build.

In Table 3, results are summarized.

Image	Conf.	Horizontal Error (m)	Vertical Error (m)	PDOP
A	LS	50,34	63,50	3,91e+04
A	RLS	42,20	59,64	2,62e+04
A	HIRLS	60,60	65,62	5,49e+04
A	WHIRLS	50,57	60,31	3,10e+04
B	LS	1,77	3,62	2,10e+02
B	RLS	1,86	3,79	2,16e+02
B	HIRLS	1,18	3,10	1,91e+02
B	WHIRLS	1,08	3,02	1,88e+02

Table 3. Camera Position Error

According with the theory, the space resection solution is more accurate when camera is close to the markers (B-image case). Indeed, when the camera is far from the object, the GCPs appear more concentrated on the image; this is confirmed by geometry analysis: in A-image case, PDOP has two orders greater than B-image so the different accuracy.

The proposed method has more accurate result for B-image while in A-image RLS has best performance.

To verify the robustness of proposed approach a blunder measurement was intentionally introduced among observation. In particular the point 2 was placed on another building corner simulating a gross error performed by the automated identification method.

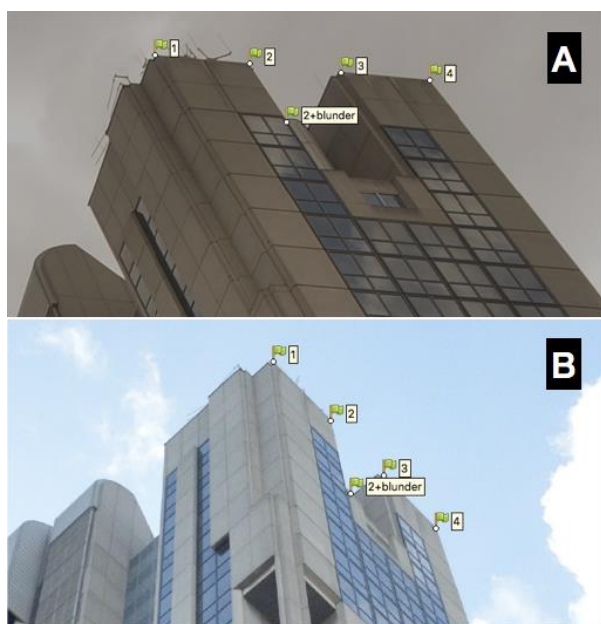


Figure 6. Blunder on point 2

With this data set, the compared configurations have the performance shown in **Errore. L'origine riferimento non è stata trovata.**

Image	Conf.	Horizontal Error (m)	Vertical Error (m)	PDOP
A	LS	53,76	84,75	3,50e+04
A	RLS	45,34	79,49	2,78e+04
A	HIRLS	49,61	62,99	3,37e+04
A	WHIRLS	44,07	59,68	2,50e+04
B	LS	4,12	3,68	1,96e+02
B	RLS	4,16	4,90	2,37e+02
B	HIRLS	2,49	4,37	2,18e+02
B	WHIRLS	2,60	4,55	2,24e+02

Table 4. Camera Position Error with blunder on marker 2

A blunder among observations of this entity (few pixels on image) has more effect on space resection performed with camera close to markers (B-image) while for A-image the camera position error has a slight deterioration. The robust proposed method has the best performance where the blunder has more effect, confirming the validity of the method.

4. CONCLUSIONS

In this paper, the performance of a non-GNSS navigation systems based on natural signal as light is studied. A robust monocular vision-based navigation system with a single camera and 3-D map is proposed. The robustness of the proposed method is verified using real data. Two collected images in an urban canyon, critical environment for GNSS positioning, are processed. The first analysis outlines are that position accuracy depends from distance between camera and points, whose coordinates are processed. In case of relative distance between camera and GCPs of about 30 meters an accuracy on position of about 50 meters occurs. This performance could be useful in the integration with other navigation system, i.e. GNSS, whose performance strongly degraded in this critical scenario.

In the first considered scenario (where no blunders are present), the proposed method provides slight improvement on closest image, while in case of outliers, intentionally considered in image coordinate of one point, the robust approach gave more promising performance.

In conclusion, the experimental results show that the vision-based navigation system guarantees an improved positioning in scenario where GNSS has problems. The robust approach gave promising results duly absorbing the effect of blunder among measurements.

In future works more experiment will be conducted; moreover, the use of a fish-eye lens will be considered, because of its ability to have in a single image a more distributed scene among the camera, improving the geometry problem.

AKNOWLEDGEMENTS

This research is supported by the University of Naples "Parthenope" and is included in the "Dipartimento di Scienze e Tecnologie (DiST)" financed project.

REFERENCES

Angrisano A., Gaglione S., and Gioia C., 2013. Performance assessment of aided global navigation satellite system for land navigation. IET Radar, Sonar Navig., vol. 7, no. 6, pp. 671–680.

- Bailey, T., and Hugh, D.W., 2006. Simultaneous localization and mapping (SLAM): Part II. *IEEE Robotics & Automation Magazine* pp 108-117.
- Box, G.E.P., 1953. Non-normality and tests on variances, *Biometrika* vol.40, pp. 318–335.
- Del Pizzo, S., & Troisi, S. 2011. Automatic orientation of image sequences in cultural heritage. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(5/W16).
- Engels, C., H. Stewnius, and D. Nistr, 2006. Bundle Adjustment Rules. *Photogrammetric Computer Vision*, 2.
- Faugeras, Olivier D., Q-T. Luong, and Stephen J. Maybank, 1992. Camera self-calibration: Theory and experiments. *European conference on computer vision*. Springer, Berlin, Heidelberg.
- Fisher, K. A. and J. F. Raquet, 2011a. Non-GPS Precision Position, Navigation, and Timing. *Air and Space Power Journal*, 26(2):24–33,
- Fisher, Kenneth A., Raquet, J. F., (2011b) Precision Position, Navigation, and Timing without the Global Positioning System. *Air & Space Power Journal*; Maxwell AFB Vol. 25, Iss. 2, pp. 24-33.
- Fraser, Clive S., 1997. Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote sensing*, pp. 149-159.
- Gaglione, S., Innac, A., Pastore Carbone, S., Troisi, S. and Angrisano, A., 2017. Robust estimation methods applied to GPS in harsh environments. In *Navigation Conference (ENC)*, 2017 European, pp. 14–25.
- Hampel, F.R., P.J. Rousseeuw, E.M. Ronchetti, and W.A. Strahel, 1986. *Robust Statistics: The approach based on influence functions*, New York: Wiley.
- Hayashi, E. and T. Kinoshita, 2009. Development of an indoor navigation system for a monocular-vision-based autonomous mobile robot. *Artificial Life and Robotics* 14(3): 324-328.
- Huber, P.J., 1964. Robust Estimation of Location Parameters, vol. 35, *The Annals of Mathematical Statistics*, Vol. 35, No. 1, pp. 73-101.
- Huber, P., 1981. *Robust Statistics*, New York: Wiley.
- Indelman V., P. Gurfil, E. Rivlin, and H. Rotstein, 2012a. Distributed vision-aided cooperative localization and navigation based on three-view geometry. *Rob. Auton. Syst.*, vol. 60, no. 6, pp. 822–840.
- Indelman V., P. Gurfil, E. Rivlin, and H. Rotstein, 2012b. Real-time vision-aided localization and navigation based on three-view geometry. *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 3, pp. 2239–2259,.
- Indelman, V., 2012c. Bundle Adjustment Without Iterative Structure Estimation and Its Application to Navigation. *Position Location and Navigation Symposium (PLANS)*.
- Knight, N.L., J. Wang, 2009. A Comparison of Outlier Detection Procedures and Robust Estimation Methods in GPS Positioning, *Journal of Navigation*, Vol. 62, pp. 699-709.
- Leick, A., 2004. *GPS Satellite Surveying*. JohnWiley and Sons, Inc.
- Li, C., Wang, J., Knight, N., and Ding, W., 2011. Vision-based Positioning with a Single Camera and 3D Maps: Accuracy and Reliability Analysis. *Journal of Global Positioning Systems*, 2011, 10, (1), p. 11
- Luhman, T., Robson, S., Kyle, S., Harley, I., 2011. *Close Range Photogrammetry, principles techniques and applications*. Whittles Publishing, 500p.
- McGlone, C., Mikhail, E. and Bethel, J., 1980. *Manual of photogrammetry*, pp.771-787.
- Mikhail, Edward M., James S. Bethel, and J. Chris McGlone, 2001. *Introduction to modern photogrammetry*. New York.
- Pierrot-Deseilligny, M., De Luca, L. and Remondino, F., 2011. Automated image-based procedures for accurate artifacts 3D modeling and orthoimage generation. *Geoinformatics FCE CTU*, pp. 291-299.
- Remondino, F., & Fraser, C. 2006. Digital camera calibration methods: considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(5), 266-272.
- Richard, H. and Zisserman, A., 2003. *Multiple view geometry in computer vision*. Cambridge university press.
- Rousseeuw, P.J., and A.M. Leroy, 1987. *Robust Regression and Outlier Detection*, New York: Wiley.
- Sabe, K., Fukuchi, M., Gutmann, J. S., Ohashi, T., Kawamoto, K., & Yoshigahara, T., 2004. Obstacle avoidance and path planning for humanoid robots using stereo vision. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on* (Vol. 1, pp. 592-597).
- Schaffrin, B., 1997. Reliability measures for correlated observations. *Journal of Surveying Engineering*, 123(3), 126–137.
- Triggs, B., P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, 2012. Bundle Adjustment: A Modern Synthesis. *2012 IEEE/ION*, 748–756.
- Troisi, S., Baiocchi, V., Del Pizzo, S., & Giannone, F., 2017. A prompt methodology to georeference complex hypogea environments. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 42.
- Wang, Chieh-Chih, et al., 2007. Simultaneous localization, mapping and moving object tracking. *The International Journal of Robotics Research* 26.9, 889-916.
- Wieser, A., and Brunner, F.K., 2002. Short static GPS sessions: robust estimation results. *GPS Solutions*, 5(3): 70–79.
- Xun, I. et al., 2011. Vision-based positioning with a single camera and 3D maps: accuracy and reliability analysis. *Journal of Global Positioning Systems*, pp. 19-29.
- Y. Yang, 1994. Robust Estimation for Dependent Observations, vol. 19, *Manuscripta Geodaetica*, pp.10-17.