Vision-based Positioning with a Single Camera and 3D Maps: Accuracy and Reliability Analysis

Xun Li, Jinling Wang, Nathan Knight, Weidong Ding
University of New South Wales
xun.li@student.unsw.edu.au

Abstract

In an effort to supplement the available satelite-based positioning technology and extend such high level positioning capability to GPS-denied environments, a method of vision-based positioning with the use of single camera and newly defined 3D maps is proposed. Besides, only natural landmarks are required in the proposed method. Absolute position and orientation information can be provided in six degree of freedom. Our work here is to address the accuracy and reliability concerns of such a vision-based navigation system. The main contribution will be the newly defined 3D map and the adoption of photogrammetric 6DOF pose estimation method to improve positioning accuracy. Dilution of Precisions (DOPs) are introduced to evaluate positioning precision within the vision-based positioning domain. Quality control strategies are also applied to detect outliers in the observation and strengthen system reliability

Keywords: *vision-based positioning; single camera; 3D map; accuracy; reliability*

_____

1. Introduction

Navigation and localization services have seen much progress in recent years. Yet most of them depend on traditional satellite-based positioning techniques and therefore can only be provided in outdoor environments (with direct GNSS visibility). In an effort to strengthen and extend the positioning capabilities to provide a more robust navigation solution, much work has been done. One of the primary concerns is to find alternative positioning techniques that could provide localization and navigation services both accurate and reliable in a GPS-denied environment. Among a great variety of studies (e.g. Priyantha et al., 2000; Want et al., 1992 Kalkusch et al., 2002), vision-based positioning is believed to be the most promising but challenging technology so far. The common approach has been using the images captured by its on-board vision sensor to match against the expectation (prior knowledge of the

navigation environment) so as to determine the position and possibly the orientation of a vision system.

To our knowledge, most of the work has come from the mobile robot community. Depending on the exploitation of one or more cameras, either map-based or mapless navigation is performed (Bonin-Font et al., 2008). Much emphasis has been placed on enabling a robot to safely and effectively navigate in an indoor environment with a high level of autonomy. However, as long as the navigation performs without failure (hitting any obstacle or unable to reach its destination), self-localization (or positioning) process is considered as satisfactory. The accuracy and reliability aspects of positioning have hardly been paid much attention, or fully investigated. This is especially true with regard to monocular vision approaches. Ohno et al. (1996) used the differences between the currently collected images and the pre-recorded image sequence to continuously estimate the robot's position and orientation shifts. While the orientation change can be obtained with relatively high accuracy, position change may not be accurately estimated. Another limitation with this approach is that it is based on the assumption that the correspondence between the current image and the reference image has always been found correctly, leaving mismatches a severe danger jeopardizing the reliability of the whole system. One year later, Ohya et al. (1997) utilized the matching of edges from the currently obtained images and the 3D edge model to achieve self-localization. A step further from the previous attempt is that they used a predetermined threshold to prevent the position error from going outrages, yet with the help of a dead-reckoning method. Visual input alone fails to provide accurate positioning results. Rivlin et al. (2003) proposed a new algorithm for image-based robot navigation applications. At the core of this idea is to generate the translation and rotation shifts in a robot movement by matching the target image with the images taken in real time. While this idea makes a good point, another contribution of their approach is that RANSAC paradigm is used to deal with outliers caused by mismatches. However, it is not without its limitations.

The algorithm is only able to provide three degrees of freedom. In several cases not enough correct matches can be found to compute the position shift. The limitation in the degrees of freedom can also be found in other approaches (Kitanov et al., 2007). Mathematical models have been developed to improve the accuracy and reliability of the system by fusing odometry information. A recent study by Hayashi and Kinoshita (2009) developed an indoor navigation system based mainly on visual input from a monocular camera and a 2D space map. Self-localization was achieved by calculating the relative position of the robot itself and straight lines recognized on both sides of the corridor, and this work didn't give explicit information on the accuracy or reliability that such method can get. It can be seen that available approaches are still far from mature to provide a robust indoor positioning and navigation solution, in a sense that a full degree of freedom should be provided with a high level of accuracy and reliability.

In an effort to move toward this direction, a method of vision-based positioning for localization and navigation purposes with the use of single camera and the newly defined 3D maps is proposed in this paper. The idea is to use classic photogrammetric mapping and positioning method in a newly devised approach to improve the positioning accuracy for vision-based navigation systems. The main contribution will be the newly defined 3D map and the adoption of photogrammetric 6DOF pose estimation method for positioning. In order to find the optimal design of this approach, great emphasis is placed on the evaluation of different factors that influence the final positioning precision and system reliability. DOP values are introduced to evaluate positioning precision within the vision-based positioning domain. Quality control strategies are also applied to detect outliers in the observation and strengthen system reliability. Besides, only natural landmarks are required to provide absolute position and orientation information.

In the next section, the methodology of our approach is introduced. In the third section, the map development procedure is discussed in detail. In the forth section the use of photogrammetric 6DOF pose estimation method for positioning and navigation is explained and justified. The fifth section introduces the use of DOP values, including basic principles and the equations derivation of the approach. The following section discusses the principles used for outlier detection in the system module. A numerical investigation is described and analyzed before the conclusion and future work are discussed.

2. Vision-based Positioning using a Single Camera and 3D maps

A method of vision-based positioning is proposed in this study with much effort to address the accuracy and reliability concerns of an indoor navigation system. By using specially built 3D maps of the navigational environment, positioning is performed based on photogrammetric method. Rather than using complete 3D models of the environment, we adopt a collection of geo-referenced images as 3D maps. Images taken by the navigation system in real time are matched with the 3D maps. The three dimensional coordinates of those matched and identified feature points are therefore transferred from the 3D maps to the current image, thus can be used as reference control points for pose estimation. We name them as pseudo ground control points (PGCPs). In the next phase, these control points are used to solve the camera positions and orientation using Photogrammetric 6DOF Pose Estimation method, which is also known as space resection. During this process, outliers from the uncertain input may affect the positioning accuracy, reliability or even lead to a positioning failure. An outlier detection module is developed to tackle this problem. With outliers being rejected, final positioning accuracy is investigated in order to find the optimal design of this approach.

The methodology involved here mainly consists of two parts. In the first step the 3D map is constructed using the geo-referenced images of the target environment. The next step is to use the 3D map built in the previous step to do positioning. The selected solution is a system with a single camera mounted on a ground vehicle. At each measuring point, an image of the environment is taken. Next, the image is processed to calculate the relative position of the camera to the known environment using a photogrammetric procedure, which forms the core of the positioning function. The data flow is shown in Figure 1.
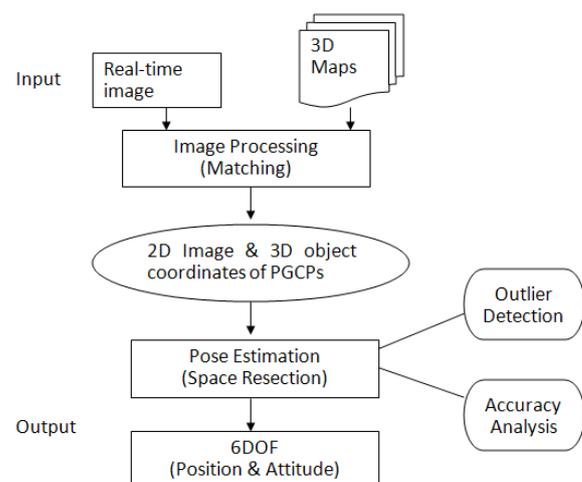

Figure 1: Data flow of the system methodology

3.   Developments of the 3D Maps

Image-based navigation is a relatively new topic in the research domain. Two typical techniques are used: optical flow-based solutions and appreance-based matching (Desouza and Kak, 2002). The first approach estimate the motion based on the analysis of a sequence of images. The second group are based on the matching of current view with stored images. Both methods are aiming at determining the position and orientation of the imaging sensor, which is assumed to have a known position and orientation relative (e.g., lever arm) to the platform that is carrying it (Hofman-Wellenhof et al., 2003).

The main difference between our approach to the available image-based navigation methods lies in the fact that the images used here are geo-referenced, which means they themselves can give absolute position information (local or global) in 3D, functioning like a sensor (e.g., GPS), and at the same time can be used as a map for any indoor localization services. Early work on the 3D maps was done by Olesk and Wang (2009).

A 3D map is defined as a sum of geo-referenced feature points with three dimensional (3D) local or global coordinates that are overlapped on images of the environment. Users of the 3D map will have the benefits of geo-referencing with 3D coordinates as well as realistic visualization. One basic function of the 3D map is for positioning and navigation. Whenever a new image is taken, it can be matched with the images stored in the 3D map database and therefore enables the user to locate its position.
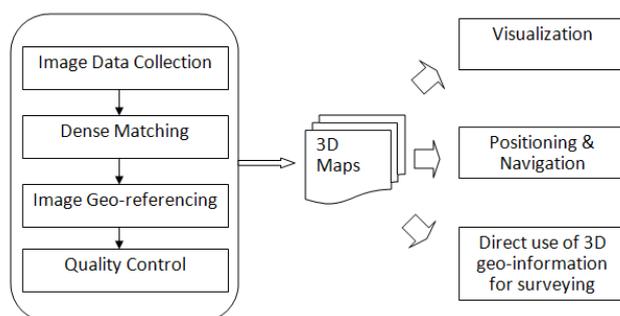
Figure 2: 3D map development

The construction of the 3D maps using geo-referenced images should be conducted with several steps. The first step is image data collection. With vision sensors mounted on a vehicle/robot, the vehicle will be guided around the navigation environment to collect image data. In the second step, dense matching of image elements over two or more images is performed to capture the 3D geometric information, establishing the correspondences between two or more images to achieve 3D information of objects in space. Suitable matching method will be adopted. Currently Lowe's SIFT (1999) method is used to extract feature points and create tie points.

In the third step, most importantly, geo-referencing of the images is performed. Currently, indirect geo-referencing, i.e. aerial triangulation method is used. After setting up and linking ground control points to the 3D environment, a bundle adjustment is utilized, and thus both the 3D environment and image position, orientations are obtained. After this, quality control for 3D mapping will be conducted in order to avoid the risk of giving useless point coordinates: during the mapping process, the effects of unmodelled errors need to be taken into consideration. Gross errors should be detected and eliminated using reliability theory for use in the bundle adjustment. The whole process is illustrated in Figure 2.

4.   Position and Orientation Determination

Photogrammetry is a classical technique to provide accurate position information. However, previous photogrammetric solutions for navigation normally based on two or more synchronized and oriented cameras to extract depth information (Luhmann, 2009). Single camera approaches have mostly been limited to the SFM (structure from motion) technique, which suffers from the problem of analyzing a complete image sequence (thus hard to process on-line and give absolute position in real time).

One method to measure position and orientation is to use a single visual sensor (e.g. camera) that mounted on the robot. The sensor looks into the navigational environment, where reference points (features) are placed and being recognized. Given a set of correspondences between known 3D reference points and their 2D positions in images, camera position and orientation can be determined (Ryberg et al., 2006). Then any object appeared in the view can also be estimated in 6DOF. It is known as pose estimation in computer vision and also called space resection in photogrammetry, which can provide highly accurate position and orientation information in 6DOF. Currently however, it has hardly been used in ground vehicle positioning and navigation. Mostly they are used in applications such as robot calibration, wheel measurement and car safety testing, where a stable reference system can be provided. The biggest problem with the method for navigation purposes lies in the fact that a reference field with the pre-designed GCPs should be available during the positioning process. This limits the positioning space to a small area, making it hard to cover long distances or be applied to non-stable/dynamic applications such as real-time positioning. In our research, this problem has been solved. Images taken in

real time can be matched with geo-referenced images map and all the matched points can be used as PGCPs for positioning.

The fundamental function model for photogrammetric 6DOF pose estimation (space resection) is called collinearity equations, which represent the geometry between projection centre, the world coordinates of an object and its image coordinates, illustrated as follows:

$$x - x_0 = -f \frac{a_1(X-X_s)+b_1(Y-Y_s)+c_1(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)}$$
$$y - y_0 = -f \frac{a_2(X-X_s)+b_2(Y-Y_s)+c_2(Z-Z_s)}{a_3(X-X_s)+b_3(Y-Y_s)+c_3(Z-Z_s)} \quad (1)$$

where x, y are the image coordinates of the object, X, Y, Z are the coordinates of the object in object frame and Xs, Ys, Zs are the coordinates of camera's perspective centre in the object frame. (x_0, y_0, f) are the camera interior parameters and (a_i, b_i, c_i) are the elements of the rotation matrix between the image and the object coordinate system. In the case of a single camera, Eq.(1) includes six unknown parameters (Xs, Ys, Zs; ω, φ, κ), in which ω, φ , κ are rotation angles around X, Y and Z axis respectively (or roll, pitch and yaw angles). The six unknowns define the 6 degrees of freedom for the vision sensor with respect to the world coordinate system. Another resection on locator points can be performed in the same sense if an unknown object appears in the view and its position needs to be estimated.

The best known method for space resection is based on a least squares solution of linearised collinearity equations. It provides the highest level of accuracy with the presence of redundant measurements, which makes it suitable for our solution with respect to the emphasis on accuracy.

5. Accuracy Measures for Vision-based Position and Orientation Determination

The applicability of one camera solutions depends strongly on the design of the reference and locator point field. Primarily the accuracy of 6DOF parameters is a function of point distribution and relative positions between the reference objects and the camera (Luhmann, 2009). In our case, relative position changes at different location, therefore the accuracy of positioning largely depends on the geometry of pseudo ground control points chosen for positioning. In order to improve position accuracy and optimize the system design, the influence of geometry change of PGCPs on the precision of position result need to be evaluated. We introduce DOP (Dilution of Precision) values. They not only enable us to observe the overall effect of geometry change on position and attitude precision, but also gain insight into the way each parameters in 6DOF is affected.

One requirement for the use of DOP is that least squares adjustment is used to estimate the final result. The least squares models are listed as follows:

$$b + v = Ax \quad (2)$$
$$D = \sigma_0^2 Q \quad (3)$$

in which Eq.(2) denotes the function model, Eq.(3) the stochastic model and σ_0 the priory standard deviation. Using this model, the covariance matrix for the estimated parameters can be obtained using Eq.(4) as:

$$C_x = \sigma_0^2 (A^T P A)^{-1} \quad (4)$$

To evaluate the impact of geometry only, the covariance of *x* will be simplified to:

$$C_x = \sigma_0^2 (A^T A)^{-1} \quad (5)$$

in which the part $(A^TA)^{-1}$ contains DOP factors in its diagonal elements. The standard deviation of each component estimated is simply the standard deviation of the inputs (σ_0) multiplied times the DOP factor. In fact, the elements in the trace of the matrix $(A^TA)^{-1}$ are functions of the geometry only, which makes it especially suitable to evaluate the effect of geometry change of control points on the final positioning accuracy. In the GPS community, DOP values are used to represent the effect of satellite geometric distribution on the accuracy of a navigation solution.

In the least squares adjustment for photogrammetric 6DOF pose estimation, for each observation (image measurement), the observation vector, design matrix and unknown vector are listed as follows:

$$b = \begin{bmatrix} x - (x) \\ y - (y) \end{bmatrix} \quad (6)$$

$$A = \begin{bmatrix} \frac{\partial x}{\partial X_s} & \frac{\partial x}{\partial Y_s} & \frac{\partial x}{\partial Z_s} & \frac{\partial x}{\partial \omega} & \frac{\partial x}{\partial \varphi} & \frac{\partial x}{\partial \kappa} \\ \frac{\partial y}{\partial X_s} & \frac{\partial y}{\partial Y_s} & \frac{\partial y}{\partial Z_s} & \frac{\partial y}{\partial \omega} & \frac{\partial y}{\partial \varphi} & \frac{\partial y}{\partial \kappa} \end{bmatrix} \quad (7)$$

$$x = [dX_s \ dY_s \ dZ_s \ d\omega \ d\varphi \ d\kappa]^T \quad (8)$$

If *n* observations are made, the complete design matrix will be:

$$A = [A_1 \ A_2 \ A_3 \ \cdots \ A_n]^T \quad (9)$$

Next the diagonal of the matrix $(A^TA)^{-1}$ is calculated as:

$$(A^TA)^{-1} = \begin{pmatrix} G_x^2 & & & & & \\ & G_y^2 & & & & \\ & & G_z^2 & & & \\ & & & G_\omega^2 & & \\ & & & & G_\varphi^2 & \\ & & & & & G_\kappa^2 \end{pmatrix} \quad (10)$$

Then we give DOP values for 6DOF, which are calculated as follows:

$$XDOP = G_x \quad YDOP = G_y \quad ZDOP = G_z \quad (11)$$

$$PDOP = \sqrt{G_x^2 + G_y^2 + G_z^2} \quad (12)$$

$$\omega DOP = G_\omega \quad \varphi DOP = G_\varphi \quad \kappa DOP = G_\kappa \quad (13)$$

$$ADOP = \sqrt{G_\omega^2 + G_\varphi^2 + G_\kappa^2} \quad (14)$$

in which the PDOP represents the Position DOP while the ADOP represents Orientation (Attitude) DOP.

6. Outlier Detection

It is well known that vision sensor has a very high input data rate, and is inherently fragile against big errors (or called observation outliers, or faults). However, FDI (fault detection and identification) has hardly been solved in the vision–based navigation domain. Outliers in our system mainly refer to gross errors coming from image coordinate measurements or PGCP coordinates. The latter occasion may be caused by mismatches, erroneous photogrammetric point determination during the process of 3D map production or in the survey of GCPs before forming the 3D map.

A classical approach to detect the outliers in geodetic observations is taken in this research. A number of contributions have been made in this area (e.g., Baarda, 1966; Baarda, 1968; Kavouras, 1982; Teunissen, 1990; Wang and Chen, 1994; 1999; Hewistion et al., 2004; Hewitson and Wang, 2006). The outlier detection mainly consists of two steps: global model test and data snooping. In the first step, global model test is carried out to check whether outlier(s) exist in the observations. It is applied on the posterior variance factor $\hat{s}_0^2$, assuming the statistic $\hat{s}_0^2/\sigma_0^2$ follows the F $_{r,\infty}$ -distribution. Given $(1-\alpha)$ confidence interval for the ratio, one tail test is recommended. If the ratio exceeds $F_\alpha(f, 1-\alpha)$, the global test fails, indicating the existence of an outlier (each time only one outlier is assumed). In the next step, data snooping is performed to identify the most likely observation contaminated by the outlier. The main idea is to consider the outlier as the mean shift of mathematical expectation of the corresponding observation. The test statistic is as follows:

$$W_i = \frac{\nabla S_i}{\sqrt{D_{\nabla S_i}}} = -\frac{e_i^T P v}{\sigma_0 \sqrt{e_i^T P Q_v P e_i}} \quad (15)$$

in which ∇S_i is the outlier in the ith observation, $D_{\nabla S_i}$ denotes its variance, P the weight matrix, Q_v the

cofactor matrix of residuals and e_i a vector of zeros with the i-th element equal to 1. W_i will follow N $(0, 1)$ distribution if no outlier exists. The largest magnitude of the value $|W_i|$ corresponds to the outlier. Given the confidence interval $(1-\alpha)$, if

$$|W_i| > N\left(0,1; 1 - \frac{\alpha}{2}\right) \quad (16)$$

then an outlier is identified.

Yet, there still can be more than one outlier in the dataset. The above precdure can be repeated to deal with multiple outliers, more discussions on this topic can be found in Knight et al. (2010). But in this study, it is assumed there is no more than one outlier. Given the power of the test β, which represents the probability of identifying a defective value as outlier, the lower bound value for non-centrality parameter δ_0 can be calculated. Using an inverse procedure of Eq.(15), the lower bound of detectable outlier $\nabla_0 S_i$, or namely the minimal detectable bias (MDB) can be calculated.

$$\nabla_0 S_i = \frac{\sigma_0 \delta_0}{\sqrt{e_i^T P Q_v P e_i}} \quad (17)$$

With a diagonal weight matrix P, Eq.(17) is simplified into:

$$\nabla_0 S_i = \frac{\delta_0}{\sqrt{r_i}} \sigma_{l_i} \quad (18)$$

A further step is to calculate the controllability $\delta'_{0,i}$, which serves as a measure of internal reliability of an adjustment system.

$$\delta'_{0,i} = \frac{\nabla_0 S_i}{\sigma_{l_i}} = \frac{\delta_0}{\sqrt{r_i}} \quad (19)$$

With several observations in the same system, the average of controllability values is calculated to give an overall level of internal reliability, with a smaller value indicating a stronger reliability.

7. Experiments

To test the methodology of the proposed system, various experiments were carried out. A calibrated CCD camera (Canon EOS4500) was used in the experiment. The focal lens is 24.7mm. The experiments were performed in the school's hallway. In the system developed for our initial test, a local object frame was adopted, where an orthogonal right-handed axis set was used with the Z

axis pointing downward. Meanwhile, the camera frame in use is illustrated in Figure 3.
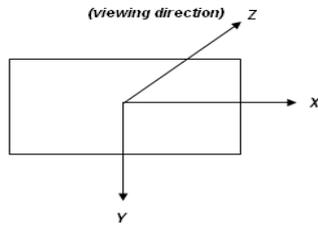

Figure. 3: Camera coordinate system

Four major experiments were carried out in this research. Experiment 7.1 was for the image geo-referencing to construct the 3D maps. Experiment 7.2 was considered to illustrate the matching process between real time images with the 3D maps. Experiment 7.3 was designed to test the outlier detection module of the positioning system. Experiment 7.4 was for the evaluation of the impact from the PGCPs on the final positioning accuracy.

7.1   Image geo-referencing to construct the 3D maps
A prototype is built at this stage of research using indirect geo-referencing, i.e. aerial triangulation method. After setting up the ground control points, a bundle adjustment is performed to estimate 3D object coordinates, image orientation parameters together with related statistical information about accuracy and reliability. The approximate values served as initial values for the unknowns in the bundle adjustment were generated using the combined intersection and resection method. The data flow of the system is shown in Fig 4.

In this initial test, three images (with overlapped area) were used and each one had a few distributed control points (more than 4) and some new (tie) points. The aim was to test the scenario of the newly defined 3D map and possible accuracy this technology could obtain. It is shown in Fig 5.
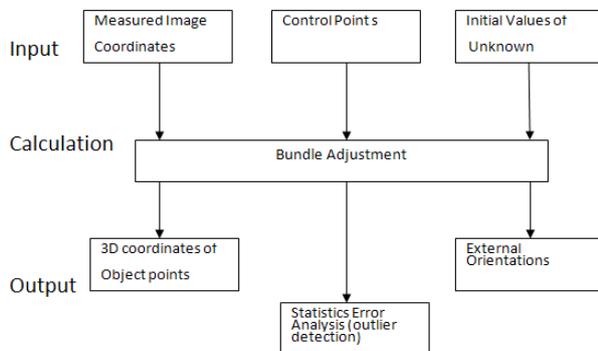

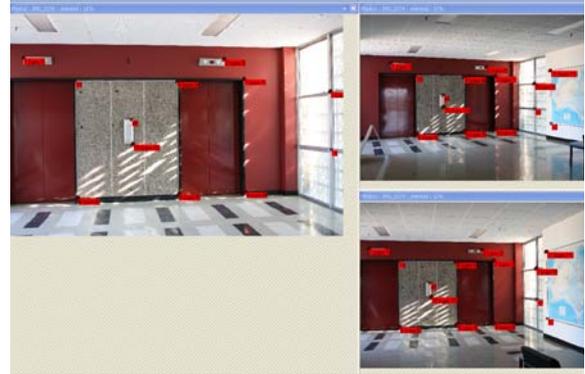Figure 4: Data flow of the prototype

Fig.5 Three images for the test

Table 1:3D Coordinates of Check Points

	Test1			Test2			Pre-Survey		
Pt No.	10	12	19	10	12	19	10	12	19
X(m)	8.4379	7.1778	10.2839	8.4252	7.1791	10.2951	8.432	7.186	10.288
Y(m)	4.5944	4.0857	0.8686	4.5872	4.0849	0.867	4.595	4.098	0.875
Z(m)	-0.8959	-0.4773	-2.0967	-0.8985	-0.4768	-2.0986	-0.896	-0.487	-2.094

Table 2: Accuracy of the 3D Map

Test No.	Average No. of GCPs	No. of Check Points	RMSE(m)			Maximum Residual (m)		
			X	Y	Z	X	Y	Z
Test1	9.3	3	0.00629	0.00801	0.00581	0.0082	0.012	0.0097
Test2	6.7	3	0.00693	0.00994	0.00662	0.0071	0.013	0.0102

Two tests were performed with different intensities of ground control points. Three check points were used to enable a rigorous evaluation of the true accuracy of the geo-referenced map. The photogrammetric determined points were compared with reference values measured by a total station. The results are shown in Tables 1 & 2. The RMSE is the root of the mean squared error between the photogrammetrically measured ground coordinates and the surveyed ground coordinates. The Maximum Residual is the maximal difference between the photogrammetrically measured ground coordinates and the surveyed ground coordinates among all the check points.

It can be observed that there is no significant difference in accuracy between axis (X, Y and Z), thoughY, the viewing direction is relatively less accurate due to the nature of photogrammetric determination. The maximum deviation found is around 0.01m, which is acceptable.

7.2 Matching with the 3D maps
Lowe 's SIFT (1999) algorithm was chosen to match the real time image with the 3D maps as its feature descriptor is invariant to scale, orientation, distortion and partially invariant to illumination changes. Initially, it was introduced to detect and describe image features for object recognition applications. We believe it is a suitable choice for our system in that when the vision

system moves in the navigation environment, features (landmarks) are observed with various viewpoints and may under different illumination. The SIFT algorithm will perform more reliable matching compared with other algorithms available.

In the first step, the SIFT features were extracted from a set of reference images and stored in a database. Basically, they are the so-called keypoints located at the maxima and minima of the difference-of Gaussian function. We extracted the SIFT features from both the 3D map database and real time image. As shown in Figure 6, 1195 keypoints were found in this image from database. They are displayed as vectors indicating scale, orientation and location.

In the next step, matching was performed based on the feature extracted. A new image was matched by individually comparing each feature from the new image to this previews database and finding the candidate matching features based on Euclidean distance of their feature vectors (Lowe, 2004). In fact, the best candidate is the nearest neighbor from the reference descriptor vector. The determination of whether it is a correct match depends on a probability measured by taking the ratio of distance from the closest neighbor to the distance of the second closest. Here we set the ratio to 0.6. Figure 7 illustrates one matching pair between the image taken in real time and the database image, in which 67 matched keypoints were produced. It should be noted that a number of mismatches were also generated during the process, which need to be dealt with at a later stage (outlier detection module).
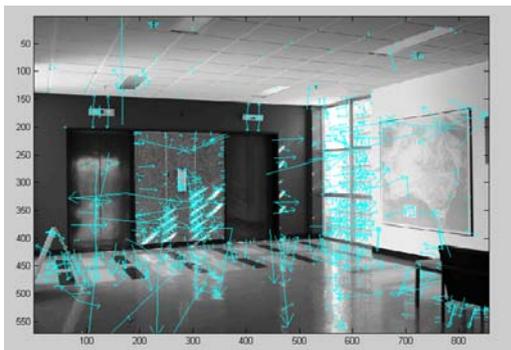

Figure 6: Extraction of SIFT feature points

Figure 7: Matching

7.3 Outlier detection for positioning

Outliers mainly come from the two uncertain inputs of position estimation function: the 2D image coordinates and the 3D object coordinates of the PGCPs. The latter occasion may be caused by mismatches, erroneous photogrammetric point determination in the 3D map developments and/or the control point surveys.
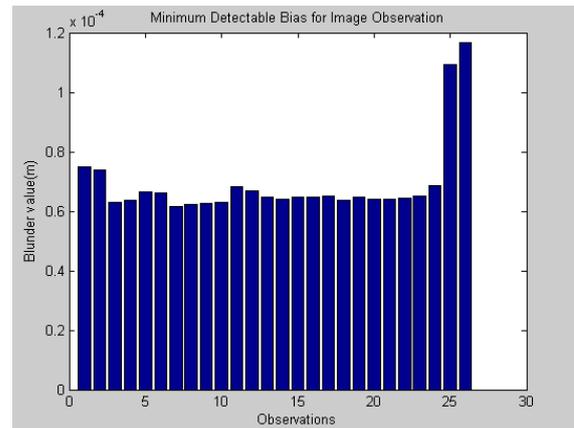

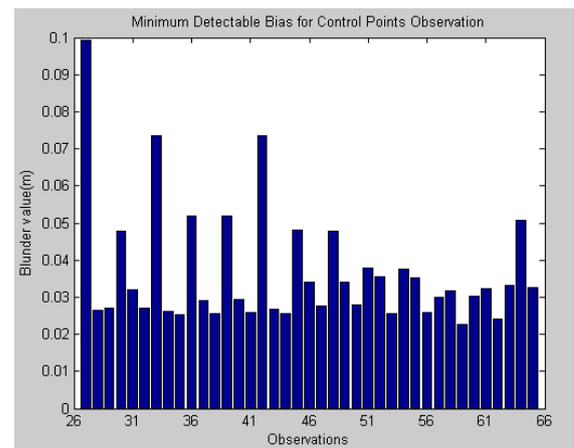Figure 8: MDB for image observations

Figure 9: MDB for PGCP observations

This experiment mainly focused on the design and testing of outlier detection module in dealing with erroneous inputs of the positioning function. The procedure was simulated using outliers intentionally inserted into a clean dataset. The image measurement noise level (priory standard deviation) was set to 0.000014 when the PGCP set to 0.00095. The posterior standard deviation of unit weight has been calculated to be 0.000014. Using the one tail global test, F $_{(20,}$ 0.95) approximately equals 1.57. F-ratio value was 1.045675, smaller than 1.57 so the global test passed with the clean data set. Figures 8 and 9 show the Minimal Detectable Bias (MDB) of each observation in the two groups of observations: the image observations and the PGCP coordinate observations. Observation No.1 to No.26 are the image observations while No.27 to No. 65 are 3D coordinates of the PGCPs. It can be seen

that these two groups of observations have different levels of MDBs, and the observations within each group do have close MDB values.

In order to simulate the impact of outliers in the image coordinates, outliers with different magnitudes were inserted intentionally into the dataset. The results were shown in Table 3 with an increasing magnitude of outliers inserted in the same image observation (Observation No. 3, x value of Point 2).

Table 3. Outler Detection for Image Measurements

	Test ID	1	2	3	4	5	6
	Outlier (mm)	0	-0.2	-0.5	-0.8	-1	-3
LS Estimation of Six Unknown Parameters (6DOF)	Xs(m)	1.0641	1.0935	1.139	1.1861	1.2183	1.5541
	Ys(m)	-2.3616	-2.4259	-2.5159	-2.5985	-2.6497	-2.9876
	Zs(m)	-1.3426	-1.3458	-1.3523	-1.3609	-1.3678	-1.4977
	ω(rad)	1.5713	1.571	1.5703	1.5694	1.5687	1.5551
	ψ(rad)	-0.0108	-0.0125	-0.015	-0.0175	-0.0192	-0.0353
	κ(rad)	2.156	2.1635	2.1745	2.1848	2.1915	2.2454
	No. 3	-0.334	-13.769	-33.924	-54.094	-67.557	-203.906
	No. 30	-0.093	-13.325	-33.169	-53.019	-66.264	200.447
W Value	No. 31	0.417	13.827	33.949	54.092	67.538	-203.873

Firstly, by using the W-statistic to locate an outlier, it was noticed that two other observations (No.30 & 31) together with observation No. 3 all produced big W values. It is noted that observation No.3 is image coordinate of PGCP Point 2-x, when No. 30 and No. 31 correspond to the X, Y value in the object space of the same point (Point 2). The three observations can be highly correlated. By studying the correlation coefficients between the W-statistics for observation pairs of No.3 and No.30, No.3 and No. 31, it was noted that the coefficient values are all close to 1 or -1, which proves the correlation is extremely strong. Secondly, it was observed that when the magnitude of outlier grows, the probability of data-snooping method successfully identifying an outlier increases. According to the result, when the magnitude is greater than 0.8, outlier is always correctly identified by data-snooping. When it is smaller than 0.8, the chance decreases. Therefore, a new outlier separability test will be developed for the module.

Table 4: Outlier Detection for 3D PGCP

	Test ID	1	2	3	4	5	6
	Outlier (mm)	0	-0.1	-0.3	-0.5	-1	-3
LS Estimation of Six Unknown Parameters (6DOF)	Xs(m)	1.0641	1.0734	1.0941	1.1179	1.1959	2.2447
	Ys(m)	-2.3616	-2.4103	-2.506	-2.5991	-2.8207	-3.26
	Zs(m)	-1.3426	-1.2854	-1.1645	-1.0334	-0.6479	2.304
	ω(rad)	1.5713	1.5779	1.5918	1.6066	1.6491	1.9629
	ψ(rad)	-0.0108	-0.01	-0.0086	-0.0071	-0.0034	0.0035
	κ(rad)	2.156	2.1635	2.1694	2.1783	2.2006	2.3039
	No.4	-1.378	14.268	45.206	75.626	148.961	371.789
W Value	No.32	1.381	-14.265	-45.205	-75.627	-148.961	-371.812

More tests were carried with outliers in the 3D coordinates of the PGCPs. The results are shown in

Table 4 with observation No.32 containing an outlier. The high correlation was found between image coordinates and the 3D coordinates of the same point. According to the results, the same conclusion can be drawn: when the magnitude of outlier grows, the probability of data-snooping method successfully identifying an outlier increases.

7.4 Geometry and Reliability Analysis

Two experiments were carried out in this section, aiming at finding the optimized solution for the selection of pseudo ground control points during positioning process. This was done by investigating the factors (mainly the geometry of pseudo control points) that influence the position accuracy.

7.4.1 Variations of the Number of PGCPs

To reveal the overall relationship between the number of PGCPs and the reliability of the system and precision of positioning, a group of tests were performed on a number of images, each tested with 15, 13,11,9,7,5,4 PGCPs respectively. One image with its results was used to show the common phenomena.

Table 5: Positioning Result in 6DOF

	Number of PGCP	15	13	11	9
LS Estimation Of Six Unknown parameters	Xs(m)	1.4531	1.4446	1.4375	1.393
	Ys(m)	0.0813	0.1066	0.1027	0.273
	Zs(m)	-1.314	-1.3165	-1.2973	-1.3053
	ω(rad)	1.5764	1.576	1.5781	1.5772
	ϕ(rad)	-0.0136	-0.0129	-0.0133	-0.0129
	κ(rad)	1.9663	1.9632	1.9634	1.944
	Number of PGCP	7	6	5	4
LS Estimation Of Six Unknown parameters	Xs(m)	1.3674	1.359	1.3407	1.3486
	Ys(m)	0.4205	0.5101	1.921	0.7081
	Zs(m)	-1.0514	-1.0869	-2.4409	-0.967
	ω(rad)	1.6048	1.601	1.45	1.615
	ϕ(rad)	-0.0189	-0.0177	0.006	-0.0161
	κ(rad)	1.9271	1.9171	1.7605	1.8957

Table 5 shows the positioning result with the use of this image. It can be seen that the estimation results of the external parameters (6DOF) tend to remain relatively stable with an increased number of PGCPs. The DOP values and the average of controllability values for each set up (e.g. 9 PGCPs) were also calculated. Figure 10 shows the variation trend of DOP values with the increase of PGCPs. Figure 11 show the average of internal controllability values with the increase of PGCPs. The three figures have further proved that the whole system is unstable with less than 13 PGCPs. The rest part of the three figures all shows a decreasing trend of the test values (DOP values and average of internal control values), which means with the increase of the number of PGCPs, the precision of positioning is increasing and the internal reliability of the system has been improved. According to the figures, it can also be observed that the increase in PGCP number has more

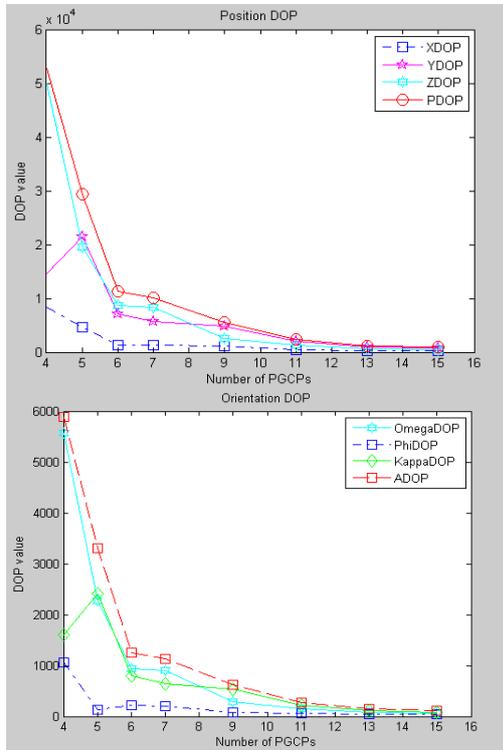impact on the precision in Z compared with the precision in X and Y.

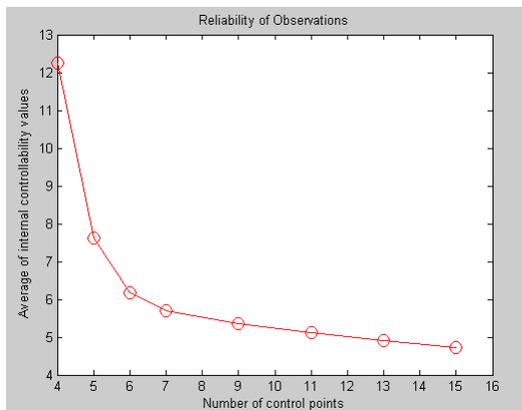Figure 10: DOP values for position and orientation

Figure 11: Internal reliability

Therefore it is concluded that PGCPs should be selected as many as possible to enable an acceptable positioning capability. When the PGCPs obtained for a particular image are not sufficient to provide a stable and relatively precise positioning results, that image for positioning should be rejected. A second image needs to be taken and matched with the 3D maps for positioning.

7.4.2 Distribution of the PGCPs
In order to investigate how the distribution of pseudo ground control points affect the positioning precision

and reliability of the system, two groups of tests were further performed on each of the real time images.

For the first group, we chose two sets of PGCPs, with one set scattered around the image and the other set centered on a small region located on the image centre. Table 6 shows the result of one image, 7 PGCPs were used for each set of this case. The estimation results of position and orientation parameters (6DOF) are close to the best results obtained previously with 15 PGCPs, which means the positioning function run successfully and the result is acceptable with both settings. It can be easily observed from DOP values that the precision of positioning is much higher with the scattered PGCPs than with the centred distribution. The internal reliability of the system has not changed much.

Table 6. Positioning Result with Scattered and Centred Distribution

		7 PGCP scattered	7 PGCP centred
	Distribution		
LS	Xs(m)	1.0952	1.0958
Estimation	Ys(m)	-2.446	-2.3501
Of	Zs(m)	-1.3226	-1.3134
Six	ω(rad)	1.5733	1.5742
Unknown	φ(rad)	-0.0115	-0.0115
parameters	κ(rad)	2.1646	2.157
	X DOP	748.105225	4441.710041
	YDOP	1889.069529	6149.407959
	ZDOP	1204.459151	4630.356573
DOP Values	PDOP	2361.983692	8887.306022
	ω DOP	137.002452	451.104166
	φ DOP	70.058549	108.223677
	κ DOP	203.636394	725.128196
	A DOP	255.236465	860.824159
Internal Reliability	Ave ControlV	5.625227	5.514999

The second one aims at investigating how the geometry change of PGCPs, especially from planar to non-planar will affect the positioning precision and system internal reliability. The tests were designed in the way that all three sets had 8 points in common and lay on the same plane. Only one point out of 9 located at different places, with the first test had the point on the same plane, the second test had the point located on a different plane and third test had the point located on the same second plane but with bigger deviation from the optical axis. The change of DOP values is shown in Figure 12.

It can be observed that a non-planar configuration of PGCPs increases the precision of the positioning result. It shows that the effect becomes more significant with the increasing offset from the optical axis. From Figure 12, it can also be observed that the precision in Z is again more affected than that of X and Y, and Omega again being the least affected among the three angle values. It is also noted from the result that internal reliability deteriorates (the average of controllability value grows). This is mainly because the points on the different planes contribute to the geometry largely, thus making it hard to be controlled. It will be difficult to

detect any outlier in this observation. In order to improve the precision of positioning and at the same time do not sacrifice system reliability, PGCPs on different planes should be selected evenly.
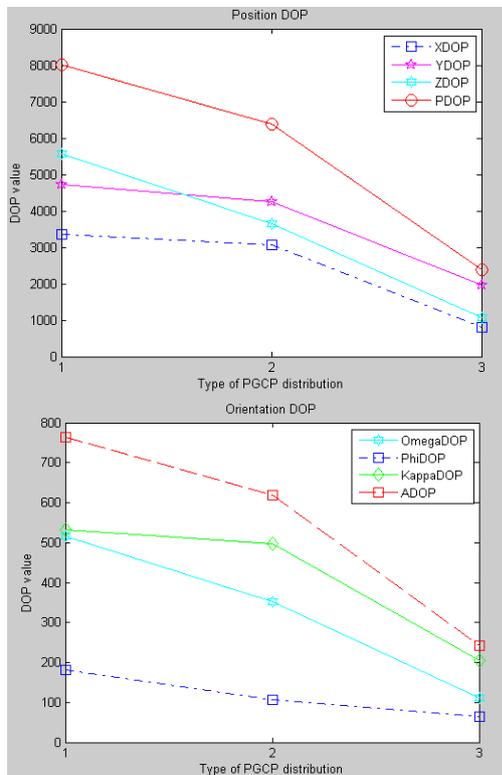

Figure 12: DOP values with three types of distribution

8. Concluding Remarks

This paper has investigated a method of vision-based positioning using a single camera and newly defined 3D maps. The aim is to develop an alternative positioning technique that can supplement satelite-based positioning in GPS-denied environment so as to achieve ubiquitous positioning when the accuracy and reliability of the positioning system can be maintained. Possible ways to improve the positioning accuracy and system reliability are discussed based on various experiments. Such techniques will have practical values in several application domains such as mobile robot navigation, emergency services, transportation, security and visitor guiding in an indoor environment.

One limitation is that the outlier detection module is not robust enough in dealing with multiple outliers as well as the separation of correlated observation containing outliers. This problem will be investigated further.

Future study is to further develop the 3D map defined and investigate the integration of vision with other sensors in order to provide a robust poisoning and navigation solution. The ultimate goal of the study is to develop alternative positioning technique that fill in the gap between outdoor and indoor localization and navigation services with respect to both accuracy and reliability concerns.

References

Baarda, W. (1966), *Statistical concepts in geodesy*. Netherlands Geodetic Commission, Publications on Geodesy, Delft, Vol. 2(4), 74pp.

Baarda, W. (1968), *A testing procedure for use in geodetic network.* Netherlands Geodetic Commission Publications On Geodesy, Delft, Vol. 2(5), pp.97.

Bonin-Font F., A. Ortiz, and G. Oliver (2008), *Visual navigation for mobile robots: a survey*. Journal of Intelligent and Robotic Systems, 53(3):263–296.

Desouza, G.N.; Kak, A.C. (2002), *Vision for mobile robot navigation: a survey*, Pattern Analysis and Machine Intelligence, IEEE Transactions on,vol.24,no.2,pp.237-267.

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.

Hewistion S. H.K. Lee & J. Wang (2004), *Localizability Analysis for GPS/Galileo Receiver Autonomous Integrity Monitoring*, Journal of Navigation, 7(2) 249-259.

Hewitson S and J. Wang (2006), *GNSS receiver autonomous integrity monitoring (RAIM) performance analysis,* GPS Solutions, vol. 10, pp. 155–170 .

Hofman-Wellenhof B., Legat K., and Wieser M (2003) *Navigation: Principles of Positioning and Guidance*. Austria: Springer-Verlag Wien New York.

Kalkusch M., T. Lidy, N. Knapp, G. Reitmayr, H. Kaufmann, and D. Schmalstieg (2002), *Structured Visual Markers for indoor pathfinding*, Pro. In IEEE Int. Workshop Augmented Reality, pp. 1-8.

Kavouras,M. (1982),*On the Detection of Outliers and the Determination of Reliability in Geodetic Networks*, Technical Report No 87. Dept. of Surveying Engineering, Univ. of New Brunswick, Fredericton, N.B., Canada.

Knight, N., Wang, J., Rizos, C., (2010). *Generalized measures of reliability for multiple outliers*, J Geod, 84, pp. 625-635.

Kitanov, A.; Bisevac, S.; Petrovic, I. (2007), *Mobile robot self-localization in complex indoor environments using monocular vision and 3D model*, Advanced intelligent mechatronics, 2007 IEEE/ASME international conference on , vol., no., pp.1-6, 4-7.

Lowe D. G. (1999), *Object Recognition from Local Scale-Invariant Features*, IEEE International Conference on Computer Vision, 2:1150.

Lowe D. G. (2004), *Distinctive Image Features from Scale-Invariant Keypoints*, International Journal of Computer Vision, 60(2):91–110.

Luhmann T (2009), *Precision potential of photogrammetric 6DOF pose estimation with a single camera*, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 64, Issue 3, Pages 275-284.

Ohno, T., A. Ohya, and S. Yuta (1996), *Autonomous navigation for mobile robots referring pre-recorded image sequence*, in Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst., Nov. pp. 672–679.

Ohya A., A. Kosaka, and A. Kak (1997), *Vision-Based Navigation of Mobile Robot with Obstacle Avoidance by Single Camera Vision and Ultrasonic Sensing*, Proc. IEEE Int'l Conf. Intelligent Robots and Systems, pp. 704-711.

Olesk, A., & Wang, J. (2009), *Geometric and error analsysis for 3D map-matching*. IGNSS Symp. 2009, Gold Coast, Australia, 1-3 December, CD-ROM procs.

Priyantha N. B., A. Chakraborth, H. Balakrishnan (2000), *The cricket location-support system*, Proc. ACM Mobicom, pp. 32-43.

Rivlin E., I. Shimshoni, and E. Smoyler (2003), *Image-based robot navigation in unknown indoor environment*. In IEEE Int. Conf. on Intell. Robots and Systems, pages 2736–2742.

Ryberg, A.; Christiansson, A.-K.; Eriksson, K. (2006), *Accuracy Investigation of a Vision Based System for Pose Measurements*, Control, Automation, Robotics and Vision, 2006. ICARCV '06. 9th International Conference on , vol., no., pp.1-6, 5-8 .

Teunissen, P.J.G. (1990), *Quality control in integrated navigation systems*, Aerospace and Electronic Systems Magazine, IEEE , vol.5, no.7, pp.35-41, Jul 1990

Want R., A. Hopper, V. Falcao, and J. Gibbons (1992), *The active badge location system*, ACM Trans. on Information Systems, Vol. 10, No. 1, pp. 91-102.

Wang J, Chen Y (1994), *On the reliability measure of observations*. Acta Geodaet et Cartograph Sin. pp 42–51 (English edition)

Wang J. & Y.Q. Chen (1994), *On the localizability of blunders in correlated coordinates of junction points in densification networks*, Australian Journal of Geodesy, Photogrammetry and Surveying, pp, 109-119.

Wang, J., Chen, Y., (1999) *Outlier detection and reliability measures for singular adjustment models*, Geomat Res Aust, 71, pp. 57-72

Biography

Xun LI received her B.E. degree in Software Engineering from Wuhan University, China, in 2009. Currently, she is working towards the Ph.D. degree at the School of Surveying & SIS, the University Of New South Wales. Her research activities are in vision-based navigation with 3D maps and GPS/INS integration. She is also interested in mobile mapping.
Email: xun.li@student.unsw.edu.au