ORIGINAL ARTICLE   **Open Access**

CrossMark

Camera auto-calibration in GPS/INS/stereo camera integrated kinematic positioning and navigation system

Nilesh S. Gopaul[*], Jianguo Wang and Baoxin Hu

Abstract

This paper presents a novel two-step camera calibration method in a GPS/INS/Stereo Camera multi-sensor kinematic positioning and navigation system. A camera auto-calibration is first performed to obtain for lens distortion parameters, up-to-scale baseline length and the relative orientation between the stereo cameras. Then, the system calibration is introduced to recover the camera lever-arms, and the bore-sight angles with respect to the IMU, and the absolute scale of the camera using the GPS/INS solution. The auto-calibration algorithm employs the three-view scale-restraint equations (SRE). In comparison with the collinearity equations (COL), it is free from landmark parameters and ground control points (GCPs). Therefore, the proposed method is computationally more efficient. The results and the comparison between the SRE and COL methods are presented using the simulated and road test data. The results show that the proposed SRE method requires less computation resources and is able to achieve the same or better accuracy level than the traditional COL.

Keywords: Camera auto-calibration, Lens distortion, Relative orientation, Lever arms, Bore-sight angles, GPS, IMU, Scale restraint equation

Introduction

The high demand for low-cost multi-sensor kinematic positioning and navigation systems as the core of direct-georeferencing technique in mobile mapping is continuously driving more research and development activities. The effective and sufficient utilization of images is among the most recent scientific research and high-tech industry development subjects. In this particular field, York University's Earth Observation Laboratory (EOL) is engaging in the study of the image-aided inertial integrated navigation as the natural continuation of its past research in the multi-sensor integrated kinematic positioning and navigation (Qian et al. 2012; Wang et al. 2015).

An image-aided inertial navigation system (IA-INS) implies that the errors of an inertial navigator are estimated via the Kalman filter using measurements derived from images. The image-based navigation algorithms, such as visual odometry (VO) (Konolige et al. 2011; Scaramuzza and Fraundorfer 2011; Gopaul et al. 2014, 2015) or visual Simultaneous Localization and Mapping

(SLAM) (Durrant-Whyte and Bailey 2006; Williams and Reid 2010; Lategahn et al. 2011; Alcantarilla et al. 2012), usually assume that a camera system is calibrated prior to its use and the calibration parameters do not change over time. The internal camera parameters (focal length, principal point and lens distortion) and the external camera parameters (baseline and relative orientation between cameras, lever-arms and bore-sight angles with respect to the inertial measurement unit (IMU)) are required to relate the image coordinates with the object coordinates in the scene. The process of estimating these parameters is referred to as the camera calibration.

The traditional camera calibration consists of capturing images containing an array of the reference targets in a laboratory, whose coordinates are accurately known (Wolf and Dewitt 2000). However, these parameters can be invalidated during in-field operations, e.g., during camera assembly/disassembly, replacement, bumps (Teller et al. 2010) or significant temperature variations. Recently many developments have focused on the in-field camera auto-calibration (or self-calibration) for image-inertial systems. An auto-calibration refers to the determination of the

* Correspondence: nileshgo@yorku.ca; nileshgopaul@gmail.com
Lassonde School of Engineering, York University, Toronto, Canada

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 2 of 15

camera parameters from a sequence of the overlapping images without necessarily setting up ground control points (GCPs) or special calibration targets. Typically, auto-calibration process is performed in a bundle adjustment (BA) (Triggs et al. 1999) or in the SLAM framework (Civera et al. 2009; Kelly and Sukhatme 2009; Kelly et al. 2011; Keivan and Sibley 2014). It involves the simultaneous estimation of the positions and orientations of the camera, the positions of the stationary landmarks, and the calibration parameters of the camera. The corresponding mathematical equation system, which models the parameters through the available measurements, is usually solved by using non-linear least-squares, the Levenberg-Marquardt algorithm or a Kalman filter. These methods however are computationally expensive due to the very large number of landmark position parameters.

Accordingly, this paper proposes a novel camera calibration method that can precisely calibrate the internal and external camera parameters with a GPS/INS/Stereo camera system exclusive of the landmark position parameters. The method applies the three-view scale-restraint equation (Bethel 2003; Ghosh 2005), with which the measurements are processed exclusively in the image space without landmark parameters. Therefore, it does not allocate large memory and computation resources. The remainder of the paper is organized as follows. Related work section overviews the related work. Then, the novel algorithm is proposed in Two-step camera calibration method section, which is followed by test results using the simulated and real data as Test results and analysis section. Conclusions section ends the paper with discussions, and conclusions.

Related work

Bender et al. (2013) presented an in-flight graph based the BA approach for system calibration between a rigidly mounted camera and an inertial navigation system. Image point features and GPS aided-INS position and orientation solution were used as measurements. Their method simultaneously computed the internal camera parameters as well as the 6-dof transformation (i.e. lever arms and relative orientation) between the two systems. However their method also required at least one GCP in-order to recover the z-component of the lever-arm. Kelly and Sukhatme (2009) proposed a camera-IMU self-calibration method within the SLAM framework implemented by an unscented Kalman filter. The lever-arms and mounting angles, the IMU gyroscope and accelerometer biases, the local gravity vector and landmarks could all be recovered from camera and IMU measurements alone. However, they assumed that the internal camera parameters were known beforehand. (Mirzaei and Roumeliotis 2008) presented a similar

tightly-coupled approach using an iterative extended Kalman filter, but, in need of known landmark position.

The methods in (Bender et al. 2013; Kelly and Sukhatme 2009; Kelly et al. 2011) implemented structure-from-motion (SfM) and contains stationary landmark parameters. SfM algorithms, which compute 3D coordinates from 2D image correspondences, have some disadvantages. The 3D Cartesian coordinates of distant objects are biased (Sibley et al. 2005) and are not well represented by Gaussian distributions (Civera et al. 2008). Similar problems arise when the baseline length between the stereo cameras and the distance between the consecutive frames are small in monocular vision (Scaramuzza and Fraundorfer 2011). Furthermore, the inclusion of landmark position in the parameter vector has two main drawbacks. First, the BA and SLAM implementation requires a good initial guess which can be difficult especially in monocular vision and when landmarks that were far away. Second, the number of landmark parameters can be very large which can result in difficult and computationally expensive estimation. Efforts to reduce the computational load were introduced in (Dang et al. 2009) where a 3D landmark position was decomposed in to 1D feature depth parameter by algebraically eliminating the x and y components using equations from the stereo pair. However, it still required the estimation of the landmark depth, a parameter not particularly useful in the calibration procedure.

Auto-calibration algorithms require a minimal constraint to define the network datum, which can be done by applying the minimum constraint, free-network adjustment, or through an explicit minimal control point (Remondino and Fraser 2006). In the free-network adjustment situation, the absolute scale of the camera system cannot be known without additional information. (Kelly et al. 2011) focused on determining the absolute scale of both the scene and the baseline in a stereo rig using GPS measurements. Their approach was similar to photogrammetric BA and the structure from motion algorithms. They could recover the baseline and relative orientation between the two cameras and lever-arms between the GPS antenna and reference camera. Similar to their previous method in (Kelly and Sukhatme 2009), they assumed that the internal camera parameters were known beforehand. Three or more overlapping image frames are required in order to estimate the camera motion on a common scale. Structure-free motion algorithms typically relied on three-view constraints (Yu et al. 2006; Indelman et al., 2013) for the same reason. The advantage is obvious. They could be exclusive of landmark position parameters and result in a more efficient algorithm.

The proposed method consists of two steps; firstly, the three-view scale-restraint equation is used to perform the free-network auto-calibration in a stereo camera
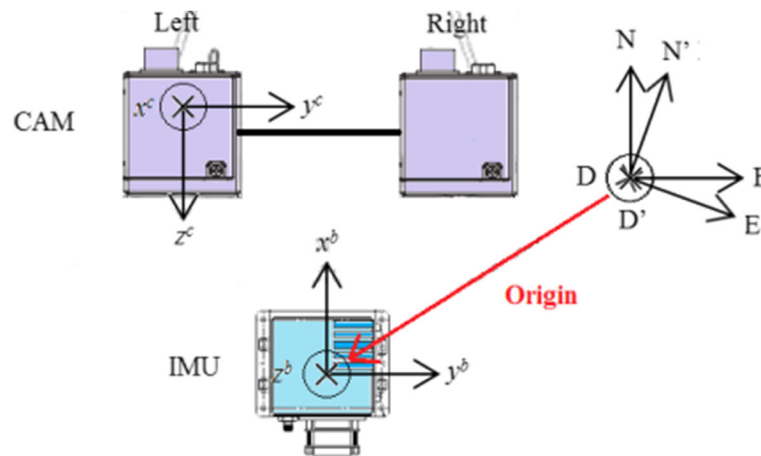
Fig. 1 Reference frame definition

system. This enables all images to operate on a common scale. And then the GPS/INS solution is applied to recover the absolute scale, as well as the bore-sight angles and the lever-arms with respect to the IMU.

Methods

In essence, a non-linear least-squares algorithm is employed to estimate the internal camera parameters, the stereo baseline and the relative orientation, the lever-arms and bore-sight angles using image measurements together with the integrated GPS aided-inertial solution. The method proceeds in two steps: (a) the stereo auto-calibration using only image measurements and then (b) the system calibration using the GPS-aided inertial integrated navigation solution (position and orientation) as external measurements. It functions under the assumptions

- The object points in the scene are stationary;

- The raw measurements from the sensors are synchronized;
- The GPS/INS blended solution has been processed;
- The GPS/INS position is referenced at the center of the IMU.

Reference frames

The following four coordinate systems are used throughout this paper (Fig. 1).

The navigation frame (n-frame) is a frame that moves with the vehicle with its origin located at a predefine point on the vehicle. Its z-axis is normal to the reference ellipsoid and points downwards while its x and y axes point towards the geodetic North and East, respectively forming a right-handed Cartesian coordinate system.

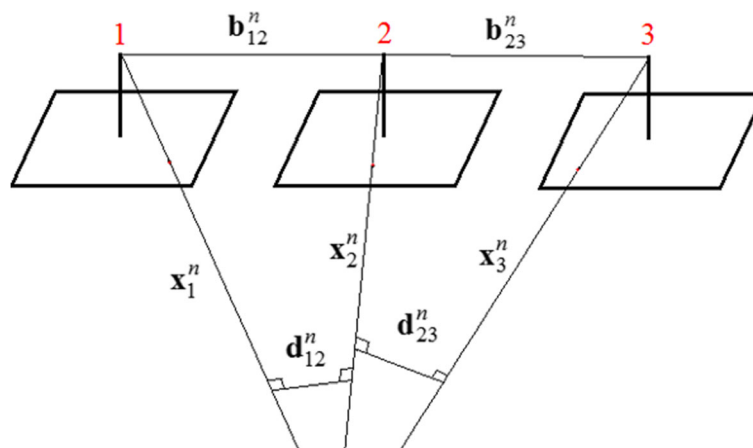The (n'-frame) has the same origin as the n-frame. Its orientation is arbitrary but fixed with respect to the n-frame.

Fig. 2 Three view geometry

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 4 of 15

The body frame (b-frame) shares the same origin with the n-frame. Its x-axis points along the vehicle's longitudinal axis and the z-axis points down while its y-axis forms a right-handed coordinate system.

The camera frame (c-frame) is a frame in which the image measurements are taken. Its origin is at the perspective center of the reference camera. Its x-axis and y-axis are parallel to the columns and rows of the charge-coupled device (CCD) sensor while its z-axis points away from the CCD sensor to form a right handed coordinate system. The camera system is assumed to be rigidly mounted on the vehicle. Hereafter, the left camera is set as the reference camera in the stereo system.

Measurement equations
Collinearity equations
The algorithmic development starts with the well-known extended collinearity equation (COL) which relates the object point position vector $\mathbf{X}_i^n$[m], its corresponding image point (x_i, y_i)[px], the camera's perspective center $\mathbf{X}^n$[m] and the direction cosine matrix (DCM)C_c^n as follows:

$$x_i = x_0 - f\left(\frac{\left[C_{c,11}^n, C_{c,21}^n, C_{c,31}^n\right] \cdot (\mathbf{X}_i^n - \mathbf{X}^n)}{\left[C_{c,13}^n, C_{c,23}^n, C_{c,33}^n\right] \cdot (\mathbf{X}_i^n - \mathbf{X}^n)}\right) + \Delta x_{d,i} + v_{x_i}$$

$$y_i = y_0 - f\left(\frac{\left[C_{c,12}^n, C_{c,22}^n, C_{c,32}^n\right] \cdot (\mathbf{X}_i^n - \mathbf{X}^n)}{\left[C_{c,13}^n, C_{c,23}^n, C_{c,33}^n\right] \cdot (\mathbf{X}_i^n - \mathbf{X}^n)}\right) + \Delta y_{d,i} + v_{y_i}$$

$$(1)$$

Where $(x_0, y_0), f$ and $\left(v_{x_i}, v_{y_i}\right)$ are the principal point, the focal length and the measurement noises, respectively. In the presence of lens distortion, the image coordinates of a point deviate from its true ones by $(\Delta x_{d,i}, \Delta y_{d,i})$, which can be modelled by (Brown 1971):

Table 1 Auto-calibration parameters (L and R denote left and right camera)

Parameter	Description
$\Delta f_L, \Delta f_R$	Correction for focal lengths [px]
$\Delta x_{L,0}, \Delta y_{L,0} \Delta x_{R,0}, \Delta y_{R,0}$	Correction for principal points [px]
$k_{L,1}, k_{L,2}, k_{L,3} k_{R,1}, k_{R,2}, k_{R,3}$	Radial lens distortion parameters [px^{-2}, px^{-4}, px^{-6}]
$\mathbf{b}_{LR}^c$	Stereo baseline vector [m]
C_{cR}^c	Right camera to left camera DCM defined by Euler angles $\boldsymbol{\theta}_{cR}^c$[deg]
$\Delta \mathbf{X}_{L,k,k-1}^{'}$	Position difference of the left camera between two consecutive frames [m]
$C_{c,k}^{n'}$	Camera to n'-frame DCM defined by Euler angles $\boldsymbol{\theta}_{c,k}^n$[deg]

$$\Delta x_{d,i} = -\Delta x_0 - \frac{\overline{x}}{f}\Delta f + \overline{x}(k_1 r^2 + k_2 r^4 + k_3 r^6 + ...)$$
$$+ p_1(r^2 + 2\overline{x}^2) + 2p_2 \overline{xy} - A_1\overline{x} + A_2\overline{y}$$

$$\Delta y_{d,i} = -\Delta y_0 - \frac{\overline{y}}{f}\Delta f + \overline{y}(k_1 r^2 + k_2 r^4 + k_3 r^6 + ...)$$
$$+ 2p_1\overline{xy} + p_2(r^2 + 2\overline{y}^2)$$

$$(2)$$

where r is the radial distance from the principal point to the image point $\left(r^2 = \overline{x}^2 + \overline{y}^2 = (x_i - x_o)^2 + (y_i - y_o)^2\right), \Delta f$ is the focal length error, $(\Delta x_0, \Delta y_0)$ is the principal point error, k_i and p_i are the coefficients of radial distortion and decentering distortion, respectively, and A_i are the affine deformation parameters. Most of the radial lens distortion is generally accounted by second term $k_2 r^4$ (Barazzetti et al. 2011). The terms with k_3 and even with k_4 term are typically included in higher-accuracy applications and wide-angle lenses. Here, the decentering distortion and affine deformation models will not be applied because they are generally very small. Furthermore, their errors will be absorbed by other terms, for example, the principal point (Fraser 2013).
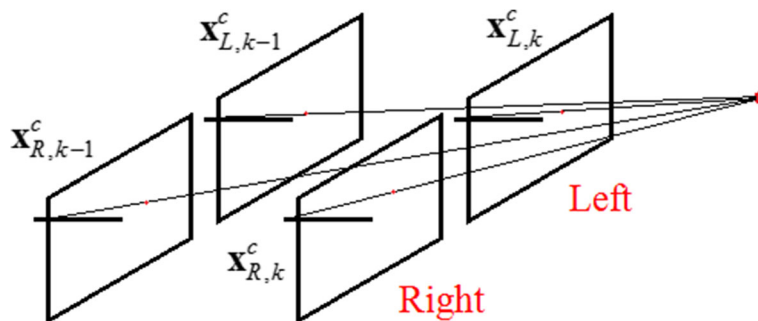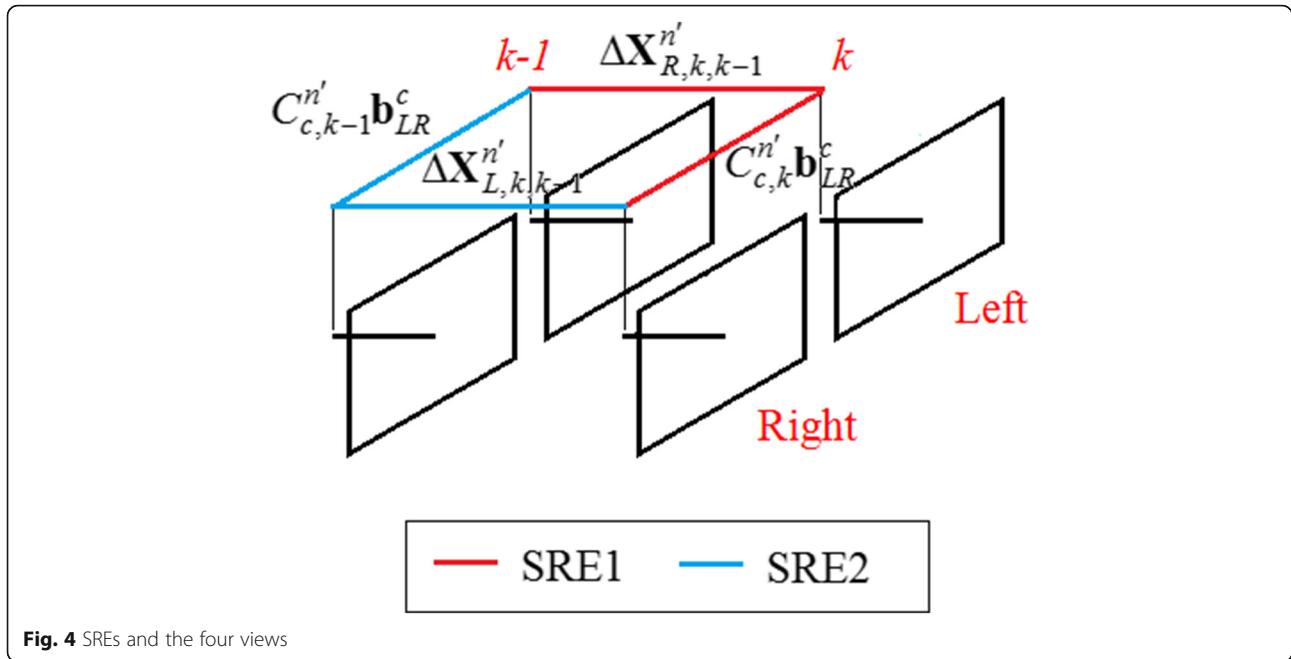

Fig. 3 Four view match

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 5 of 15

Fig. 4 SREs and the four views

Scale restraint equation

The scale restraint equation (SRE) is a robust three-view constraint that forces the images to operate on one common scale. Let's consider image point vectors $\mathbf{x}_1^n, \mathbf{x}_2^n$ and $\mathbf{x}_3^n$ (Fig. 2):

where

$$\mathbf{x}_{i=1,2,3}^n = C_{c,i}^n \begin{pmatrix} x_i - x_o - \Delta x_{d,i} - v_{x_i} \\ y_i - y_o - \Delta y_{d,i} - v_{y_i} \\ -f \end{pmatrix} \tag{3}$$

and $\mathbf{b}_{12}^n$, $\mathbf{b}_{23}^n$ are the baselines [m] between images 1 and 2, and between images 2 and 3, respectively. If image 1 is relatively oriented to image 2, and image 2 is relatively oriented to image 3, there is no guarantee that image 1 and image 3 are also relatively oriented to each other (Bethel 2003). As a result, three vectors $\mathbf{x}_1^n, \mathbf{x}_2^n$, and $\mathbf{x}_3^n$ will fail to intersect at a common point due the scale variations of the three views. The 'mismatch' vector $\mathbf{d}_{12}^n$ is perpendicular to both of $\mathbf{x}_1^n$ and $\mathbf{x}_2^n$, and is computed as $\mathbf{d}_{12}^n = \mathbf{x}_1^n \times \mathbf{x}_2^n$. Similarly the 'mismatch' vector $\mathbf{d}_{23}^n = \mathbf{x}_2^n \times \mathbf{x}_3^n$. Midway along the vectors $\mathbf{d}_{12}^n$ and $\mathbf{d}_{23}^n$ is the point where the two adjacent vectors are the closest. With the vectors in Fig. 2, one can give

$$\begin{aligned} k_1 \mathbf{x}_1^n - k_2 \mathbf{x}_2^n + k_{12} \mathbf{d}_{12}^n &= \mathbf{b}_{12}^n \\ k_2' \mathbf{x}_2^n - k_3' \mathbf{x}_3^n + k_{23}' \mathbf{d}_{23}^n &= \mathbf{b}_{23}^n \end{aligned} \tag{4}$$

where $k_1, k_2, k_{12}, k_2', k_3'$ and k_{23}' are unknown scalars with unique values as follows (Ghosh 2005)

$$\begin{aligned} k_1 &= \frac{\mathbf{b}_{12}^n \cdot \mathbf{d}_{12}^n \times \mathbf{x}_2^n}{\mathbf{x}_1^n \cdot \mathbf{d}_{12}^n \times \mathbf{x}_2^n} & k_2' &= \frac{\mathbf{b}_{23}^n \cdot \mathbf{d}_{23}^n \times \mathbf{x}_3^n}{\mathbf{x}_2^n \cdot \mathbf{d}_{23}^n \times \mathbf{x}_3^n} \\ k_2 &= \frac{\mathbf{x}_1^n \cdot \mathbf{d}_{12}^n \times \mathbf{b}_{12}^n}{\mathbf{x}_1^n \cdot \mathbf{d}_{12}^n \times \mathbf{x}_2^n} & k_3' &= \frac{\mathbf{x}_2^n \cdot \mathbf{d}_{23}^n \times \mathbf{b}_{23}^n}{\mathbf{x}_2^n \cdot \mathbf{d}_{23}^n \times \mathbf{x}_3^n} \\ k_{12} &= \frac{\mathbf{x}_1^n \cdot \mathbf{b}_{12}^n \times \mathbf{x}_2^n}{\mathbf{x}_1^n \cdot \mathbf{d}_{12}^n \times \mathbf{x}_2^n} & k_{23}' &= \frac{\mathbf{x}_2^n \cdot \mathbf{b}_{23}^n \times \mathbf{x}_3^n}{\mathbf{x}_2^n \cdot \mathbf{d}_{23}^n \times \mathbf{x}_3^n} \end{aligned} \tag{5}$$

Their analysis in detail can be found in Appendix A. In order for all the vectors to intersect at the same point, $k_2 + k_2'$ must equal to zero (Ghosh 2005), i.e.

$$\frac{\mathbf{x}_1^n \cdot \mathbf{d}_{12}^n \times \mathbf{b}_{12}^n}{\mathbf{x}_1^n \cdot \mathbf{d}_{12}^n \times \mathbf{x}_2^n} - \frac{\mathbf{x}_2^n \cdot \mathbf{d}_{23}^n \times \mathbf{b}_{23}^n}{\mathbf{x}_2^n \cdot \mathbf{d}_{23}^n \times \mathbf{x}_3^n} = 0 \tag{6}$$

where in $\mathbf{d}_{12}^n = \mathbf{x}_1^n \times \mathbf{x}_2^n$ and $\mathbf{d}_{23}^n = \mathbf{x}_2^n \times \mathbf{x}_3^n$. Equation (6) is the scale restraint equation that forces the independent scale factors for the common ray between the stereo pair 1–2 and stereo pair 2–3 to be equal (Bethel 2003).

Table 2 Dimension of the parameter vector (COL vs. SRE)

	COL	SRE
Number of image frames	n_x	n_x
Number of observed landmarks	n_{lm}	n_{lm}
Focal length, principal point	2x3	2x3
Lens distortion (k_1, k_2, k_3)	2x3	2x3
Stereo baseline and relative orientation	$2 + 3$	$2 + 3$
Camera position and orientation	$6 (n_x - 1)$	$6 (n_x - 1)$
Landmark parameters	$3n_{lm}$	0
Total	$11 + 6n_x + 3n_{lm}$	$11 + 6n_x$

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3
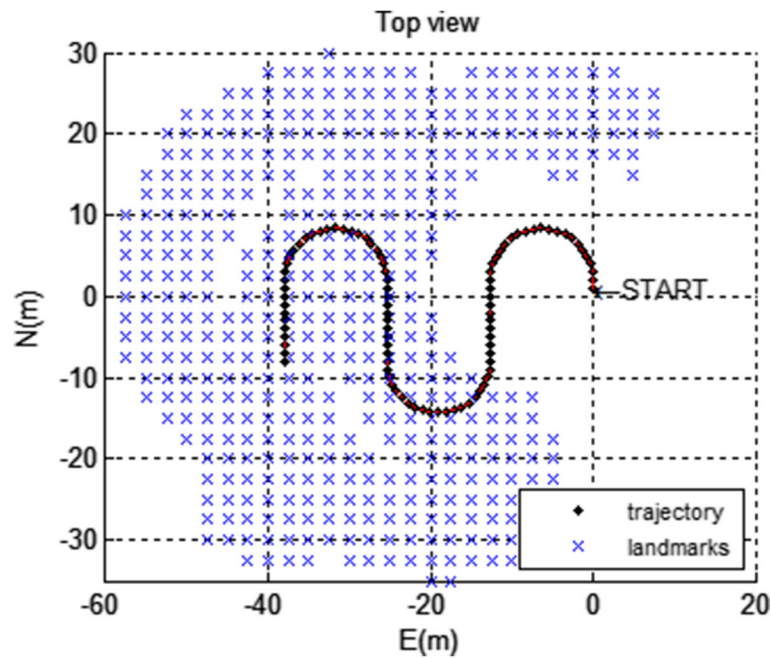
Page 6 of 15

Fig. 6 The 2D trajectory and landmarks

This equation is mainly used in successive relative orientation of image pairs and scale transfer.

Camera auto-calibration

Table 1 lists the relevant calibration parameters. At a given epoch k, the algorithm uses point features matched from four views which are stereo pairs from two consecutive epochs, i.e., $\mathbf{x}_{L,k}^c, \mathbf{x}_{R,k}^c, \mathbf{x}_{L,k-1}^c$ and $\mathbf{x}_{R,k-1}^c$ as depicted in Fig. 3.

Point features can be extracted using the Harris corner detector (Harris and Stephens 1988) and matching can be performed using the Sum of Absolute Differences (SAD) in an 11×11 window. To improve the matching results between stereo pairs, the search is performed along the epipolar lines (Bin Rais et al. 2003). Furthermore, to improve the matching between the consecutive frames, the locations of the features in the current frame can be predicted from the previous frame using the inertial sensors (Veth et al. 2006).
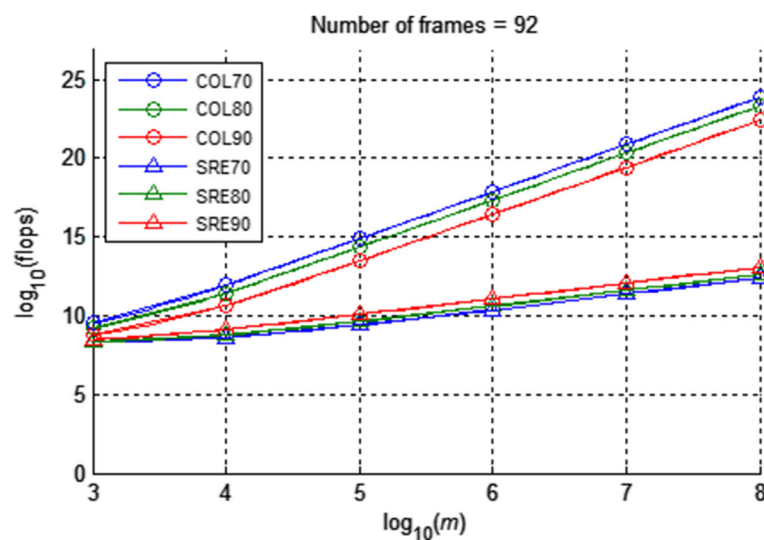

Fig. 5 Flops vs. number of stereo points (*m*)

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3
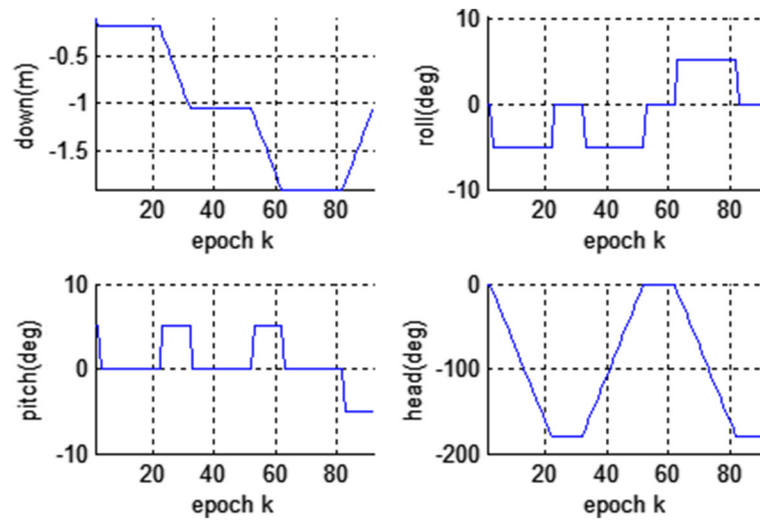
Page 7 of 15

Fig. 7 The height, roll, pitch and heading profiles

The matched points are constrained by two SREs. The first SRE relates $\mathbf{x}_{R,k}^c, \mathbf{x}_{L,k}^c$ and $\mathbf{x}_{L,k-1}^c$ while the second one relates $\mathbf{x}_{L,k-1}^c, \mathbf{x}_{R,k-1}^c$ and $\mathbf{x}_{R,k}^c$ (Fig. 4).

The equations can be expressed as

$$\frac{\mathbf{x}_{R,k}^{n'} \bullet \left(\mathbf{x}_{R,k}^{n'} \times \mathbf{x}_{L,k}^{n'} \right) \times C_{c,k}^{n'} \, \mathbf{b}_{LR}^c}{\mathbf{x}_{R,k}^{n'} \bullet \left(\mathbf{x}_{R,k}^{n'} \times \mathbf{x}_{L,k}^{n'} \right) \times \mathbf{x}_{L,k}^{n'}}$$

$$-\frac{\mathbf{x}_{L,k}^{n'} \bullet \left(\mathbf{x}_{L,k}^{n'} \times \mathbf{x}_{L,k-1}^{n'} \right) \times \Delta\mathbf{X}_{L,k,k-1}^{n'}}{\mathbf{x}_{L,k}^{n'} \bullet \left(\mathbf{x}_{L,k}^{n'} \times \mathbf{x}_{L,k-1}^{n'} \right) \times \mathbf{x}_{L,k-1}^{n'}} = 0$$

$$(7)$$

$$\frac{\mathbf{x}_{L,k-1}^{n'} \bullet \left(\mathbf{x}_{L,k-1}^{n'} \times \mathbf{x}_{R,k-1}^{n'} \right) \times C_{c,k-1}^{n'} \mathbf{b}_{LR}^c}{\mathbf{x}_{L,k-1}^{n'} \bullet \left(\mathbf{x}_{L,k-1}^{n'} \times \mathbf{x}_{R,k-1}^{n'} \right) \times \mathbf{x}_{R,k-1}^{n'}}$$

$$-\frac{\mathbf{x}_{R,k-1}^{n'} \bullet \left(\mathbf{x}_{R,k-1}^{n'} \times \mathbf{x}_{R,k}^{n'} \right) \times \Delta\mathbf{X}_{R,k,k-1}^{n'}}{\mathbf{x}_{R,k-1}^{n'} \bullet \left(\mathbf{x}_{R,k-1}^{n'} \times \mathbf{x}_{R,k}^{n'} \right) \times \mathbf{x}_{R,k}^{n'}} = 0$$

$$(8)$$

where $\mathbf{x}_{L,k}^{n'} = C_{c,k}^{n'} \mathbf{x}_{L,k}^c, \mathbf{x}_{R,k}^{n'} = C_{c,k}^{n'} C_{cR}^c \mathbf{x}_{R,k}^{cR}$ and $\Delta\mathbf{X}_{R,k,k-1}^{n'} = \Delta\mathbf{X}_{L,k,k-1}^{n'} + \left(C_{c,k}^{n'} - C_{c,k-1}^{n'} \right) \mathbf{b}_{LR}^c$. In auto-calibration, the orientation of the camera with respect to the n-frame is

1. Simulate trajectory (i.e. position & attitude)
2. Simulate landmark position (i.e. map)
3. For each position and attitude
   a. Find landmarks in view
   b. Find landmarks with ranges [15,25] m
   c. For each landmark
      i. Get image coordinates
      ii. Add lens distortion
      iii. Add measurement noise (0.0±0.5px)

Fig. 8 The simulation algorithm

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3
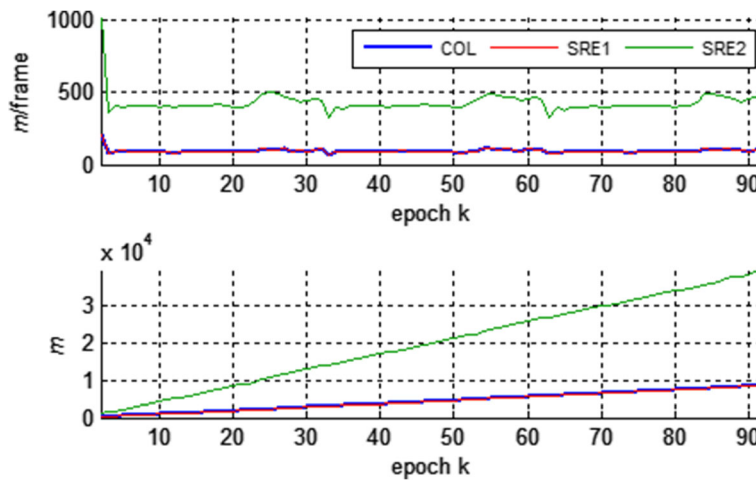
Page 8 of 15

Fig. 9 Number of stereo points per frame (*top*) and number of cumulative stereo points (*bottom*)

not required and can be put aside. At this point the *global* frame is set to the n'-frame. In order to obtain a free-network adjustment, one component of the baseline vector $\mathbf{b}_{LR}^c$ must be free (i.e. 2 dof) and one of the $\boldsymbol{\theta}_c^{n'}$ orientation parameters must be fixed (ideally $\boldsymbol{\theta}_{c,k=1}^{n'} = \mathbf{0}$). This fixes both the orientation and scale of the system. Note that $\Delta\mathbf{X}_{L,k}^{n'}$ and $C_{c,k}^{n'}$ contain transport rate error in this formulation. Under the assumption that the calibration area is within a few hundred meters, this error effect is negligibly small.

Camera system calibration

The camera system calibration determines the lever-arms $\mathbf{la}_L^b$, the absolute scale s_c of the camera and the bore-sight angle vector $\boldsymbol{\theta}_c^b$. The relationship between the GPS/INS and the left camera can be expressed as the following 7-parameter transformation:

$$\mathbf{X}_{GPSINS,k}^n = s_c C_{n'}^n \mathbf{X}_{L,k}^{n'} - C_{b,GPSINS,k}^n \mathbf{la}_L^b \qquad (9)$$

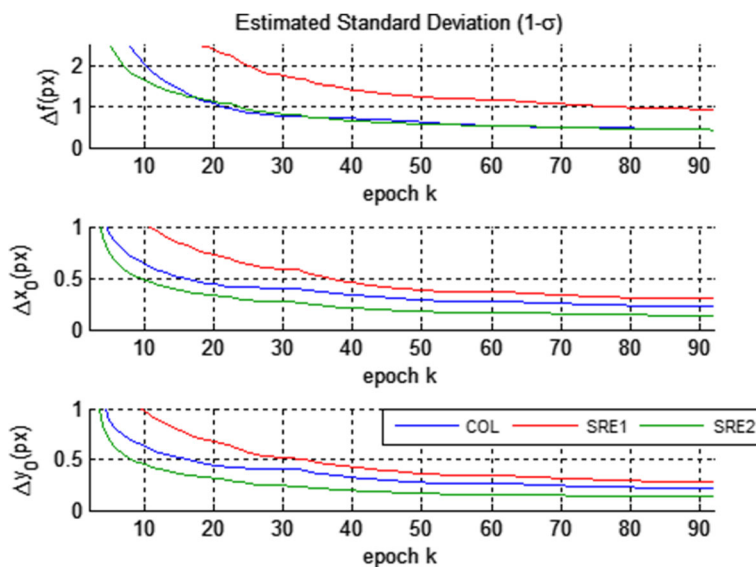where in $C_{b,GPSINS,k}^n$ is the GPS/INS DCM as the function of attitude angles (roll, pitch and heading) and $C_{n'}^n$

Fig. 10 Estimated standard deviation of the focal length error and principal point error

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3
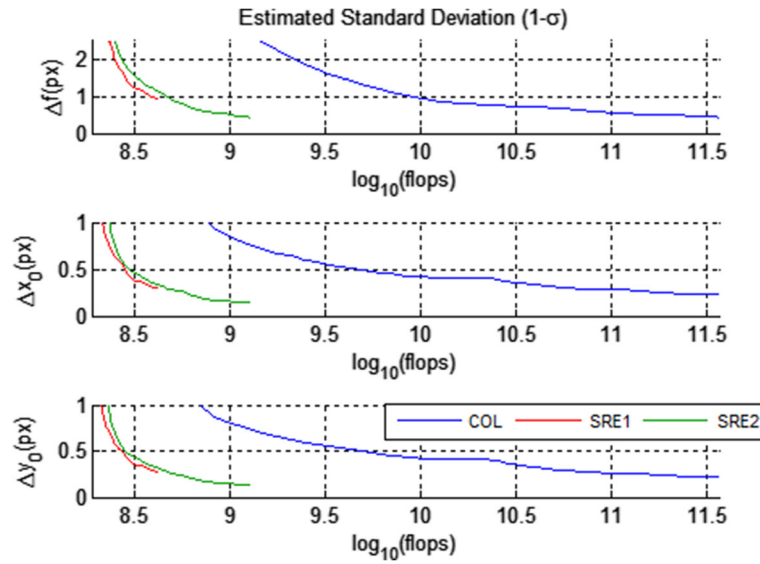
Page 9 of 15

Fig. 11 Estimated standard deviation of the focal length error and principal point vs. flops

is determined by the orientation of the camera system with respect to the n-frame at the first epoch

$$C_{n'}^n = C_{c,1}^n = C_{b,GPSINS,1}^n C_c^b \qquad (10)$$

where C_c^b is the DCM from the camera to body represented by the bore-sight angle vector $\boldsymbol{\theta}_c^b$. Differencing (8) between epochs $k-1$ and k gives

$$\Delta\mathbf{X}_{GPSINS,k}{}^n = s_c C_n^n \Delta\mathbf{X}_{L,k}^{n'} - \left(C_{b,GPSINS,k}^n - C_{b,GPSINS,k-1}^n\right)\mathbf{la}_L^b \qquad (11)$$

The relationship between $C_{b,GPSINS,k}^n$ and the camera DCM $C_{c,k}^{n'}$ (from Eqs. (7) and (8)) can be written as

$$C_{b,GPSINS,k}^n C_c^b = C_{n'}^n C_{c,k}^{n'} \qquad (12)$$

Equations (11) and (12) equate the GPS/INS information $\left(\Delta\mathbf{X}_{GPSINS,k}^n, C_{b,GPSINS,k}^n\right)$ and the auto-calibration estimates $\Delta\mathbf{X}_{L,k}^{n'}$ and $C_{c,k}^{n'}$. All of seven parameters can be solved by using the least-squares.

Computation complexity of COL and SRE

This section compares the number of parameters and the number of floating point operations (flops) between COL and SRE auto-calibration algorithms.

Table 2 shows the number of parameters with COL and SRE auto-calibration system with n_x stereo image frames and n_{lm} visible landmarks. One component of the stereo baseline is left free (i.e., only two stereo

Table 3 Left camera lens distortion parameters

Parameter	True value	COL	SRE1	SRE2
Δf (px)	−2	−1.825	−2.713	−2.318
		±0.438	±0.875	±0.417
Δx_0 (px)	2.5	2.393	2.665	2.432
		±0.230	±0.309	±0.142
Δy_0 (px)	−3	−3.232	−2.851	−3.061
		±0.219	±0.266	±0.122
k_1 (px^{-2})	5.0e^{-7}	5.03e^{-07}	5.07e^{-07}	5.06e^{-07}
		±6.47e^{-09}	±6.97e^{-09}	±3.21e^{-09}
k_2 (px^{-4})	4.0e^{-13}	7.16e^{-14}	6.49e^{-13}	3.92e^{-13}
		±1.03e^{-13}	±1.04e^{-13}	±4.75e^{-14}
k_3 (px^{-6})	4.5e^{-19}	1.78e^{-18}	4.89e^{-18}	3.72e^{-18}
		±4.86e^{-19}	±4.89e^{-19}	±2.20e^{-19}

Table 4 Right camera lens distortion parameters

Parameter	True value	COL	SRE1	SRE2
Δf (px)	+2	2.082	2.534	2.078
		±0.441	±0.884	±0.421
Δx_0 (px)	−2	−2.003	−2.250	−2.029
		±0.226	±0.319	±0.146
Δy_0 (px)	1	0.391	1.217	1.233
		±0.217	±0.243	±0.112
k_1 (px^{-2})	5.0e^{-07}	5.08e^{-07}	5.04e^{-07}	5.06e^{-07}
		±6.72e^{-09}	±8.02e^{-09}	±3.76e^{-09}
k_2 (px^{-4})	4.0e^{-13}	2.85e^{-13}	2.38e^{-13}	3.60e^{-13}
		±1.08e^{-13}	±1.28e^{-13}	±6.03e^{-14}
k_3 (px^{-6})	4.5e^{-19}	9.01e^{-19}	5.89e^{-18}	5.82e^{-18}
		±5.21e^{-19}	±6.22e^{-19}	±2.93e^{-19}

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 10 of 15

baseline parameters). The number of the camera position and orientation parameters is 6 $(n_x{-}1)$ in total because the first camera position and orientation is fixed (practically they are set to zero). The advantage of employing SRE is to have the number of the estimated parameters far less than the one with COL.

The flop count is the total number of textbook multiplication and addition operations required to obtain a least squares (LS) solution. The factors taken into account in the analysis are the number of the matched stereo points (i.e. number of measurements), the number of the image frames, the number of the landmarks in view and the percentage overlap between consecutive frames. The percentage overlap encompasses camera rate, the velocity and the angular rate of the camera. Furthermore, COL employs the LS algorithm in the explicit form (i.e., $z = h(x)$) to estimate the parameters while the SRE uses implicit LS (i.e., $h(x, z) = 0$)); where x is the parameter vector, z is the measurement vector and h(.) is the functional model. The flop counts between the two will be different under a given number of measurements and parameters.

In order to simplify the analysis, the number of image frames is kept constant (here set to 92, the same number used in the test results). Furthermore, the number of measurements, landmarks and the overlapping percentage is assumed to have the following predictive relationship

$$m' = \frac{m}{n_{lm}} = \frac{1}{1-p} \tag{13}$$

where m' is the average number of the matched stereo pairs per landmark, m is the total number of the matched stereo pairs and p is the average overlapping percentage. For instance, if $p = 75\%$ then $m' = 4$. This means on average a landmark is viewed on four images.

Table 5 Relative orientation of right camera w.r.t left camera

Parameter	True value	COL	SRE1	SRE2
$b^c_{LR,x}$(m)	0.01	0.011	0.007	0.007
		±0.001	±0.002	±0.001
$b^c_{LR,y}$(m) [a]	0.65	0.650	0.650	0.650
$b^c_{LR,y}$(m)	−0.1	−0.013	−0.013	−0.012
		±0.004	±0.004	±0.002
$\theta^c_{CR,x}$(deg)	−0.25	−0.257	−0.266	−0.260
		±0.007	±0.008	±0.004
$\theta^c_{CR,y}$(deg)	0.5	0.504	0.500	0.494
		±0.008	±0.014	±0.006
$\theta^c_{CR,z}$(deg)	0	0.002	−0.003	−0.002
		±0.002	±0.002	±0.001

[a]Free parameter

Table 6 Number of points and parameters

	COL	SRE1	SRE2
Number of stereo points	17,074	17,074	77,945
Number of parameters	8417	563	563
$\log_{10}$ (flops)	12.4	8.8	9.4

By keeping the average percentage overlap constant, the number of landmarks in the system can be predicted with a given number of stereo points. The number of measurements and parameters in the LS are now known and therefore the flop count can be predicted. Figure 5 shows the number of flops vs. number of stereo points required in COL and SRE with overlapping of 70, 80 and 90%.

As expected, the plot shows that COL uses more flops than SRE. As percentage overlap increases, the number of flops in COL decreases because the number of the matched stereo pairs per landmark becomes larger. Therefore, given the same number of measurements, the number of the landmark parameters becomes smaller. As percentage overlap increases, the number of flops in SRE increases because more matrix inversion operations are needed in the implicit LS algorithm. The accuracy analysis is presented in Auto-calibration results section.

Results and discussion

In this section test results from the simulated, laboratory and real data are presented. Simulations were performed to validate the proposed SRE auto-calibration algorithm and to show how its performance (both computation and accuracy) in comparison with the one from the COL auto-calibration method. Finally results from land vehicle data are presented.

Results from the simulated data

The simulations were conducted to compare the performance of COL and SRE auto-calibration algorithms based on a typical land vehicle trajectory (i.e. large horizontal motion and heading variation). Figure 6 shows the vehicle's trajectory and the landmarks. The vehicle's height and attitude profiles are given in Fig. 7.

The camera resolution and the field of view (FOV) were set to 640 × 480 pixels and 50° (equivalent to 686.2 pixels), respectively. The baseline between the two cameras is 0.65m long. Figure 8 describes the simulation parameters. The number of the epochs is 92. To simulate

Table 7 IMU440CA Specification

Angular Rate	Bias Stability [deg/h]	<10.0
	Angle Random Walk [deg/√hr]	<4.5
Acceleration	Bias Stability [mg]	<1.0
	Velocity Random Walk [m/s/√hr]	<1.0

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3
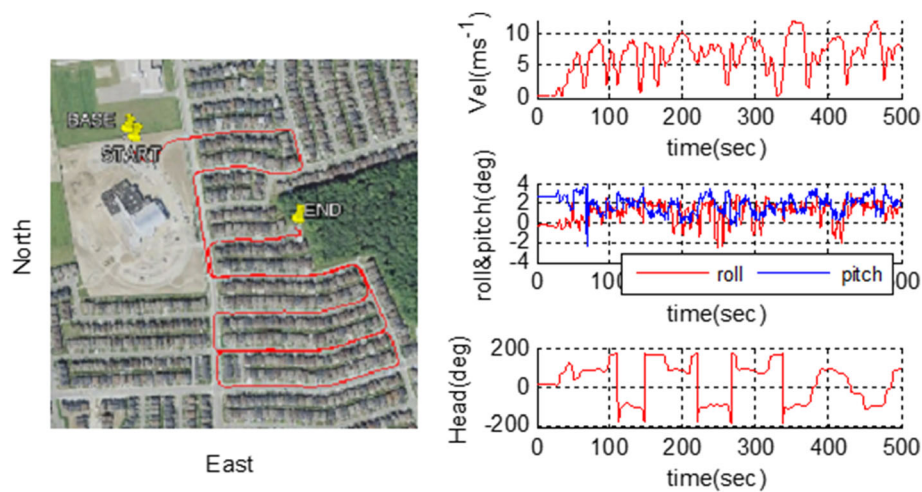
Page 11 of 15

Fig. 12 The 2D overview of the trajectory (*left*), the velocity and attitude profiles (*right*)

the urban scenario, the landmarks between the ranges 15 and 25 were selected to be in view of the camera. Furthermore, the measurement noise was set to zero mean Gaussian noise with the standard deviation of 0.5px. The initial value for each camera calibration parameter was set to zero except the baseline component $b_{LR.y}^{c}$ as the free parameter and equal to 0.65m.

Auto-calibration accuracy analysis

The accuracy analysis on COL and SRE algorithms is presented in this section. Auto-calibration results from one COL and two SREs are presented (i.e. SRE1 and SRE2). The estimates from COL and SRE1 were obtained using the same number of the measurements, i.e., 17,074 stereo points (m) whilst SRE2 used 4.5 times more measurements. The average percentage overlap for all three cases was 74%. Figure 9 shows the number of stereo points per frame and the number of accumulated stereo points.

Figure 10 shows the estimated standard deviation of the focal length error and principal point obtained from the three estimates w.r.t. epoch. The results showed that with a given number of measurements, SRE1 performed worse than COL. However SRE2 required 4.5 times more to be equivalent or better than COL. Figure 11 shows the estimated standard deviation w.r.t. the number of flops. The results showed that both SRE1 and SRE2 required less computation resources to achieve the same level of the accuracy as COL.

Auto-calibration results

This section presents the final results from COL, SRE1 and SRE2. Tables 3, 4 and 5 show the true values, estimates and their standard deviations for the left, right and relative camera orientation calibration parameters, respectively. The results showed that the accuracy of the focal length error from COL and SRE2 were similar and better than SRE1. SRE2 estimated the best principal point error, but SRE1 gave the worst. The radial distortion coefficients
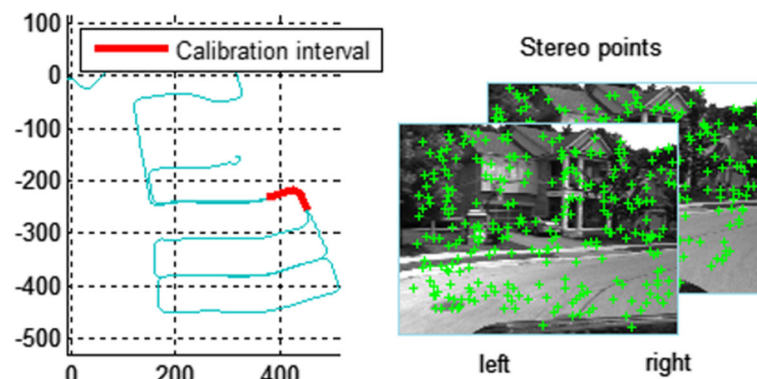

Fig. 13 The calibration period (*left*). A stereo pair with matched points (*right*)

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 12 of 15

from COL and SRE1 were similar. However, SRE2 delivered the best coefficients.

The relative orientation parameters from SRE2 were the best whilst the ones from COL and SRE1 were similar to each other. Table 6 shows the total number of the used stereo points, the number of parameters and the flop count for each of the auto-calibration algorithm. COL and SRE1 employed the same number of measurements, as SRE2 used 4.5 times more measurements. SRE1 and SRE2 estimated the same number parameters, but COL estimated 7854 more parameters. Even though SRE2 processed more measurements than COL, it still used 1000 times less flops since less number of parameters were estimated. Furthermore the accuracy is higher since more measurements were employed.

Results from road test data

Test results from both the camera auto-calibration and system calibration using road test data are presented in this section. The road data were collected by the land vehicle navigation system developed at the Earth Observation Laboratory of York University (Qian et al. 2012) with two newly integrated cameras. The system consists of two NovAtel OEM GPS receivers, one Crossbow IMU440CA and the two PointGrey Flea3 cameras as described in the lab test section. Two GPS receivers provided the absolute heading measurements and a third GPS receiver was used as the base station so that RTK level GPS positioning accuracy was achieved. The lever-arm vectors of the GPS receivers and cameras with respect to the IMU unit were measured beforehand at the accuracy of 0.5cm. The observation rates were set to 1.0, 100 and 7.5 Hz for GPs, IMU and cameras, respectively. The IMU specification is in Table 7.

The data was collected in Vaughan, Ontario and was 518 s long. Figure 12 shows the top view of the vehicle's trajectory, the velocity profile and the attitude profiles.

Table 8 Left camera lens distortion parameters

Parameter	COL	SRE1	SRE2
Δf (px)	−1.651	−1.815	−1.785
	±0.468	±0.961	±0.458
Δx_0 (px)	0.049	0.157	0.178
	±0.134	±0.177	±0.112
Δy_0 (px)	−0.245	−0.262	−0.155
	±0.756	±0.819	±0.452
k_1 (px^{-2})	−3.63e^{-07}	−3.59e^{-07}	−3.46e^{-07}
	±7.42e^{-09}	±7.51e^{-09}	±3.73e^{-09}
k_2 (px^{-4})	−7.45e^{-12}	−8.61e^{-13}	−6.61e^{-13}
	±1.03e^{-13}	±1.02e^{-13}	±5.80e^{-14}
k_3 (px^{-6})	8.02e^{-18}	5.97e^{-18}	6.64e^{-18}
	±3.45e^{-19}	±4.81e^{-19}	±3.54e^{-19}

Table 9 Right camera lens distortion parameters

Parameter	COL	SRE1	SRE2
Δf (px)	1.141	1.244	1.050
	±0.468	±0.968	±0.442
Δx_0 (px)	1.652	1.577	1.601
	±0.132	±0.174	±0.106
Δy_0 (px)	1.451	1.407	1.305
	±0.725	±0.780	±0.345
k_1 (px^{-2})	−3.34e^{-7}	−3.82e^{-07}	−3.62e^{-07}
	±7.45e^{-09}	±7.71e^{-09}	±3.84e^{-09}
k_2 (px^{-4})	−1.09e^{-12}	−4.65e^{-13}	−4.46e^{-13}
	±9.84e^{-14}	±1.04e^{-13}	±7.14e^{-14}
k_3 (px^{-6})	7.26e^{-18}	4.32e^{-18}	6.54e^{-18}
	±4.42e^{-19}	±4.83e^{-19}	±3.72e^{-19}

One hundred images in the interval shown in Fig. 13 (left) were used to test the camera calibration algorithms. This section of the trajectory was chosen because the vehicle dynamics favor lever-arm estimation. Furthermore, the images were highly textured which could be important in detecting point features. Figure 13 (right) shows the 53rd stereo image with the matched feature points. Similar to the tests performed in Results from road test data section, the calibration results from one COL and two SREs are presented (i.e. still as SRE1 and SRE2). COL and SRE1 estimates were obtained using the same number of stereo points whilst SRE2 used 5.1 times more measurements. The average percentage overlap for all three cases is 73%.

Auto-calibration results

The estimated lens distortion parameters together with their standard deviations of the left and right cameras are shown in Tables 8 and 9, respectively. The results showed the focal length error estimates from COL and SRE2 were similar, but SRE1 performed the worst. SRE2 estimated the best principal point error, followed by COL and SRE1. Noticeably, Δx_0 is better estimated than Δy_0 because the distribution of the points on the images in the x component was more varied and thus benefited the Δx_0 estimation. The results showed that the coefficients k_1 and k_2 accounted for most of the radial

Table 10 Relative orientation of right camera w.r.t left camera

Parameter	COL	SRE1	SRE2
$b_{LR,x}^c$(m)	0.002 ± 0.001	0.004 ± 0.001	0.003 ± 0.001
$b_{LR,y}^c$(m)	0.65	0.65	0.65
$b_{LR,y}^c$(m)	−0.002 ± 0.001	−0.001 ± 0.001	−0.001 ± 0.001
$\theta_{cR,x}^c$(deg)	0.250 ± 0.010	0.247 ± 0.014	0.251 ± 0.009
$\theta_{cR,y}^c$(deg)	−0.211 ± 0.003	−0.220 ± 0.004	−0.220 ± 0.003
$\theta_{cR,z}^c$(deg)	−0.093 ± 0.002	−0.092 ± 0.002	−0.091 ± 0.002

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 13 of 15

Table 11 Number of points, parameters and flops

	COL	SRE1	SRE2
Number of stereo points	26,564	26,564	109,652
Number of parameters	9260	617	617
$\log_{10}$ (flops)	13.0	9.0	9.5

distortion. Their standard deviations were the lowest from SRE2, while they were similar in COL and SRE1.

Table 10 shows the estimated relative orientation parameters and their standard deviations. The results showed that the estimated baseline vectors were similar. The angles in the y and z components were also similar, whereas the x component was similar from COL and SRE2. The worst came from SRE1. Table 11 shows the total number of the used stereo points, number of parameters and the flop count for each of the auto-calibration algorithms. COL used 3162 times more flops than SRE2. The most important finding in the tests was that despite the fact SRE2 used more measurements; it outperformed COL both in terms of accuracy and computational efficiency.

System calibration results

Table 12 shows the estimated lever-arms, absolute scale factor and bore-sight angles of the stereo camera system. The results showed that the lever-arms from COL and SRE2 were comparable, whereas the one from SRE1 was the worst. SRE2 estimated the best bore-sight angles followed by COL and SRE1. The absolute scale factors were similar in all three calibration results.

The standard deviation of the lever-arm in the z-component was almost three times larger than the ones in x and y components as there was little variation in the vehicle's pitch angle and the accuracy of the GPS position was worse vertically than horizontally. The lever-arms were measured beforehand using a measuring tape at the accuracy of 0.1cm. Table 13 shows the difference between the estimated and the measured lever-arms, and the standard deviations. The differences were within two

Table 12 Lever-arm, Scale and Bore-sight

Parameter	COL	SRE1	SRE2
$la^b_{L,x}$(m)	-0.037 ± 0.016	-0.025 ± 0.026	-0.065 ± 0.013
$la^b_{L,y}$(m)	0.352 ± 0.019	0.334 ± 0.031	0.331 ± 0.018
$la^b_{L,z}$(m)	-0.018 ± 0.047	-0.167 ± 0.090	-0.093 ± 0.040
s_c	0.983 ± 0.001	0.989 ± 0.001	0.981 ± 0.001
$\theta^b_{c,x}$(deg)	90.445 ± 0.010	90.721 ± 0.013	90.654 ± 0.008
$\theta^b_{c,y}$(deg)	-0.195 ± 0.014	-0.161 ± 0.029	-0.174 ± 0.011
$\theta^b_{c,z}$(deg)	-90.401 ± 0.031	-90.365 ± 0.054	-90.452 ± 0.024

Table 13 Difference between the estimated lever-arm and the measured lever arm components

Measured lever-arm (m)	COL (m)	SRE1 (m)	SRE2 (m)
-0.060	0.023 ± 0.016	0.035 ± 0.026	-0.005 ± 0.013
0.325	0.027 ± 0.019	0.009 ± 0.031	0.006 ± 0.018
-0.050	0.032 ± 0.047	-0.117 ± 0.090	-0.043 ± 0.040

times the standard deviations (95%) which showed that the lever-arms were correctly estimated.

Conclusions

This paper presented a novel two-step camera calibration method in a GPS/INS/Stereo camera integrated kinematic positioning and navigation system. The first step performs the camera auto-calibration for a stereo system by employing two scale-restraint equations to constrain the matched features from two consecutive stereo pairs. The lens distortion parameters, the up-to-scale baseline length and the relative orientation between the two cameras are estimated using the least-squares method. The second step performs system calibration where the auto-calibration estimates are fused with the blended GPS/INS solution to recover the camera lever-arms, the absolute scale of the camera and the bore-sight angles. The main advantage of the proposed novel method lies that it is free from landmark parameters and results in computation and memory savings. There are two main drawbacks in employing the scale-restraint equation over the collinearity equations for stereo auto-calibration. Firstly the accuracy cannot be increased by performing loop closures when the same scene is revisited. Secondly the scale-restraint equation is highly non-linear and therefore the LS estimator can diverge if a good approximation of the parameters is not available.

The results from the simulated and real road test data were presented and showed that the proposed auto-calibration method requires less computation resources to achieve equal or better accuracy than applying the traditional collinearity equations despite the fact it using more measurements.

Appendix A

From the principles of vector analysis, any four vectors **a**, **b**, **c** and **d** in three-dimensional space can be related to each other through

$$(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = (\mathbf{a}.\mathbf{c} \times \mathbf{d})\mathbf{b} - (\mathbf{b}.\mathbf{c} \times \mathbf{d})\mathbf{a}$$
$$= (\mathbf{a}.\mathbf{b} \times \mathbf{d})\mathbf{c} - (\mathbf{a}.\mathbf{b} \times \mathbf{c})\mathbf{d}$$

This implies,

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 14 of 15

$$(\mathbf{a}.\mathbf{b} \times \mathbf{d})\mathbf{c} = (\mathbf{a}.\mathbf{c} \times \mathbf{d})\mathbf{b} - (\mathbf{b}.\mathbf{c} \times \mathbf{d})\mathbf{a} + (\mathbf{a}.\mathbf{b} \times \mathbf{c})\mathbf{d}$$

$$\mathbf{c} = \frac{(\mathbf{a}.\mathbf{c} \times \mathbf{d})}{(\mathbf{a}.\mathbf{b} \times \mathbf{d})}\mathbf{b} - \frac{(\mathbf{b}.\mathbf{c} \times \mathbf{d})}{(\mathbf{a}.\mathbf{b} \times \mathbf{d})}\mathbf{a} + \frac{(\mathbf{a}.\mathbf{b} \times \mathbf{c})}{(\mathbf{a}.\mathbf{b} \times \mathbf{d})}\mathbf{d}$$

$$\mathbf{c} = b\mathbf{b} - a\mathbf{a} + d\mathbf{d}$$

Where the scalar multipliers a, b and d are

$$a = \frac{(\mathbf{b}.\mathbf{c} \times \mathbf{d})}{(\mathbf{a}.\mathbf{b} \times \mathbf{d})}$$

$$b = \frac{(\mathbf{a}.\mathbf{c} \times \mathbf{d})}{(\mathbf{a}.\mathbf{b} \times \mathbf{d})}$$

$$d = \frac{(\mathbf{a}.\mathbf{b} \times \mathbf{c})}{(\mathbf{a}.\mathbf{b} \times \mathbf{d})}$$

Authors' contributions
NG carried out the camera auto-calibration and system calibration research. NG conceived, designed, implemented and tested the algorithms presented in this paper. JW and NG collected the GPS, IMU and image data used for the tests. NG drafted the manuscript. JW reviewed and edited the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

References
Alcantarilla PF, Yebes JJ, Almazàn J, Bergasa LM (2012) On combining visual SLAM and dense scene flow to increase the robustness of localization and mapping in dynamic environments. In: Proceedings at the 2012 IEEE International Conference on Robotics and Automation, Saint Paul, Minnesota, USA, 14–18 May 2012:1290–1297. doi:10.1109/ICRA.2012.6224690.

Barazzetti L, Mussio L, Remondino F, Scaioni M (2011) Targetless Camera Calibration, International Archives of the Photogrammetry. In: ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVIII-5/W16, 2011. pp 335–342. doi:10.5194/isprsarchives-XXXVIII-5-W16-335-2011.

Bender D, Schikora M, Sturm J, Cremers D (2013) Graph-based bundle adjustment for INS-camera calibration. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W2, 2013UAV-g2013, Rostock. doi:10.5194/isprsarchives-XL-1-W2-39-2013.

Bethel JS (2003) Photogrammetry and Remote Sensing. In: Chen WF, Liew JYR (ed) The Civil Engineering Handbook, 2nd Edition, Chapter 56. CRC Press, Boca Raton, London, New York, Washington D.C, p 2170.

Brown DC (1971) Close-Range Camera Calibration. Photogramm Eng 37(8):855–866

Bin Rais N, Khan HA, Kamran F, Jamal H (2003) A new algorithm of stereo matching using epipolar geometry. In: Proceedings at the Multi Topic Conference, 2003. INMIC 2003. 7th International, Islamabad, Pakistan, 9–9 December 2003:21–24. doi:10.1109/INMIC.2003.1416609.

Civera J, Davison AJ, Montiel JMM (2008) Inverse depth parametrization for monocular SLAM. IEEE Trans Robot 24(5):932–945. doi:10.1109/TRO.2008.2003276

Civera J, Bueno DR, Davison AJ, Montiel JMM (2009) Camera self-calibration for sequential Bayesian structure from motion. In: Proceedings at the IEEE International Conference on Robotics and Automation, ICRA2009, Kobe. doi:10.1109/ROBOT.2009.5152719.

Dang T, Hoffmann C, Stiller C (2009) Continuous stereo self-calibration by camera parameter tracking. IEEE Trans Image Process 18(7):536–1550. doi:10.1109/TIP.2009.2017824.

Durrant-Whyte H, Bailey T (2006) Simultaneous localization and mapping (SLAM): part I the essential algorithms. IEEE Robot Autom Mag 13(2):99–110. doi:10.1109/MRA.2006.1638022

Fraser CS (2013) Automatic camera calibration in close-range photogrammetry. Photogramm Eng Remote Sens 79(4):381–388

Ghosh SK (2005) Fundamentals of Computational Photogrammetry. Concept Publishing Company, New Delhi, p 117-119

Gopaul NS, Wang JG, Hu B (2014) Discrete EKF with pairwise time-correlated measurement noise for image-aided inertial integrated navigation. ISPRS Ann Photogramm, Remote Sens Spat Inf Sci 2(2):61–66. doi:10.5194/isprsannals-II-2-61-2014.

Gopaul NS, Wang JG, Hu B (2015) Multi-frame Visual Odometry in Image-Aided Inertial Navigation System. J. Sun at al. (eds) China Satellite Navigation Conference (CSNC) 2015 proceedings. Lecture Notes Electr Eng 342(3):649–658. doi:10.1007/978-3-66-46632-2_57.

Harris C, Stephens M (1988) A combined corner and edge detection. In: Proceedings of the Fourth Alvey Vision Conference (1998)., pp 147–151

Indelman V, Melim A, Dellaert F (2013) Incremental Light Bundle Adjustment for Robotics Navigation. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Tokyo. doi:10.1109/IROS.2013.6696615.

Kelly J, Sukhatme GS (2009) Visual-inertial simultaneous localization, mapping and sensor-to-sensor self-calibration. In: Proceedings at the Computational Intelligence in Robotics and Automation (CIRA), 2009 IEEE International Symposium, Daejeon, South Korea, 15–18 December 2009: 360–368 doi:10.1109/CIRA.2009.5423178.

Kelly J, Matthies LH, Sukhatme GS (2011) Simultaneous mapping and stereo extrinsic parameter calibration using GPS measurement. In: Proceedings at the Robotics and Automation (ICRA), 2011 IEEE International Conference, Shanghai, China, 9–13 May 2011:279–286. doi:10.1109/ICRA.2011.5980443.

Keivan N, Sibley G (2014) Constant-time Monocular Self-Calibration. In: Proceedings at the IEEE International Conference on Robotics and Biomimetics (ROBIO) 2014. Bali, Indonesia, 5–10 December 2014:1590–1595. doi:10.1109/ROBIO.2014.7090561.

Konolige K, Agrawal M, Solà J (2011) Large scale visual odometry for rough terrain, In: The International Journal of Robotics Research 66:201-212. doi:10.1007/978-3-642-14743-2_18.

Lategahn H, Geiger A, Kitt B (2011) Visual SLAM for autonomous ground vehicles. In: Proceedings of the Robotics and Automation (ICRA), 2011 IEEE International Conference, Shanghai, China, 9–13 May 2011:1732–1737. doi:10.1109/ICRA.2011.5979711.

Mirzaei FM, Roumeliotis SI (2008) A Kalman Filter-Based Algorithm for IMU-Camera Calibration: Observability Analysis and Performance Evaluation. IEEE Trans Robot 24(5):1143–1156. doi:10.1109/TRO.2008.2004486.

Qian K, Wang JG, Gopaul NS, Hu B (2012) Low Cost Multi-sensor Kinematic Positioning and Navigation System with Linux/RTAI. J Sens Actuator Netw 1(3):166–182

Remondino F, Fraser C (2006) Digital Camera Calibration Methods: Considerations and Comparisons. The International Archives of the Photogrammetry and Remote Sensing Commission V Symposium. Image Engineering and Vision Metrology. IAPRS 36(5):266–272, Dresden 25–27 September 2006

Scaramuzza D, Fraundorfer F (2011) Visual Odometry Part I&II: The First 30 Years and Fundamentals. IEEE Robot Autom Mag 18(4):80–92

Sibley G, Matthies L, Sukhatme G (2005) Bias Reduction and Filter Convergence for Long Range Stereo. In: 12th International Symposium of Robotics Research, San Francisco 28:285-294. doi:10.1007/978-3-540-48113-3_26.

Teller S, Koch O, Walter MR, Huang AS (2010) Ground robot navigation using uncalibrated cameras. In: Proceedings at the Robotics and Automation (ICRA), 2010 IEEE International Conference, Anchorage, USA, 3–7 May 2010: 2423–2430. doi:10.1109/ROBOT.2010.5509325.

Triggs N, McLauchlan P, Hartley R, Fitzgibbon A (1999) Bundle Adjustment—A Modern Synthesis. In: Proceedings at the International Workshop on Vision Algorithms ICCV 99. Springer-Verlag, pp 298–372. doi:10.1007/3-540-44480-7_21.

Veth M, Raquet J, Pachter M (2006) Stochastic constraints for efficient image correspondence search. IEEE Trans Aerosp Electron Syst 42(3):973–982. doi:10.1109/TAES.2006.4439212.

Wang JG, Qian K, Hu B (2015) An Unconventional Full Tightly-Coupled Multi-Sensor Integration for Kinematic Positioning and Navigation, J. Sun at al. (eds) China Satellite Navigation Conference (CSNC) 2015 proceedings. Lecture Notes Electr Eng 342(3):753–765. doi:10.1007/978-3-662-46632-2_65.

Williams B, Reid I (2010) On Combining Visual SLAM and Visual Odometry. In: Proceedings at the Robotics and Automation (ICRA), 2010 IEEE International

Gopaul *et al. The Journal of Global Positioning Systems* (2016) 14:3

Page 15 of 15

Conference, Anchorage, USA, 3–7 May 2010: 3494–3500. doi:10.1109/ROBOT.2010.5509248.

Wolf PR, Dewitt BA (2000) Elements of Photogrammetry with Applications in GIS, 3rd edn.. ISBN 0-07-292454-3

Yu YK, Wong KH, Chang MMY, Or SH (2006) Recursive Camera-Motion Estimation With the Trifocal Tensor. IEEE Trans Syst, Man, Cybern, Part B (Cybern) 36(5): 1081–1090. doi:10.1109/TSMCB.2006.874133.