

## Article

# Extrinsic Parameter Calibration for Line Scanning Cameras on Ground Vehicles with Navigation Systems Using a Calibration Pattern

Alexander Wendel \*  and James Underwood 

The Australian Centre for Field Robotics (ACFR), Department of Aerospace, Mechanical and Mechatronic Engineering (AMME), The University of Sydney, Sydney, NSW 2006, Australia; j.underwood@acfr.usyd.edu.au

\* Correspondence: a.wendel@acfr.usyd.edu.au; Tel.: +61-2-9114-0895

Received: 1 September 2017; Accepted: 25 October 2017; Published: 30 October 2017

**Abstract:** Line scanning cameras, which capture only a single line of pixels, have been increasingly used in ground based mobile or robotic platforms. In applications where it is advantageous to directly georeference the camera data to world coordinates, an accurate estimate of the camera's 6D pose is required. This paper focuses on the common case where a mobile platform is equipped with a rigidly mounted line scanning camera, whose pose is unknown, and a navigation system providing vehicle body pose estimates. We propose a novel method that estimates the camera's pose relative to the navigation system. The approach involves imaging and manually labelling a calibration pattern with distinctly identifiable points, triangulating these points from camera and navigation system data and reprojecting them in order to compute a likelihood, which is maximised to estimate the 6D camera pose. Additionally, a Markov Chain Monte Carlo (MCMC) algorithm is used to estimate the uncertainty of the offset. Tested on two different platforms, the method was able to estimate the pose to within 0.06 m/1.05° and 0.18 m/2.39°. We also propose several approaches to displaying and interpreting the 6D results in a human readable way.

**Keywords:** line scan cameras; extrinsic calibration; camera pose; navigation system; GPS; ground vehicles; georeferencing

## 1. Introduction

Line scanning (also 1D or linear) cameras, which produce a single line of pixels for each exposure, have been used widely in areas such as remote sensing [1,2] and industrial inspection [3–5]. While 2D frame cameras offer the benefit of imaging a larger scene with each exposure, linescan cameras allow capturing of images at higher frame rates or spatial resolution [6]. One specific but common example is hyperspectral line scanning cameras, which provide both high spatial and spectral resolution. Many applications require accurate and direct determination of the real world coordinates of line scan image data, also known as georeferencing or mapping. This requires precise calibration of the sensor's intrinsic (e.g., focal length and principal point) and extrinsic parameters (i.e., camera pose with respect to the vehicle body frame). In the remote sensing literature, determination of extrinsic parameters is known as lever arm (translation) and boresight (orientation) alignment. More recently line scanning cameras have also been studied for low altitude unmanned aerial vehicle (UAV) and mobile ground based applications [7–10], but there are fewer studies addressing the extrinsic calibration requirements that closer proximity to the scene implies. Requirements include obtaining a 6 degree of freedom (DOF) extrinsic parameter solution including translation, which has a greater influence on mapping when proximal; avoiding ground control points (GCPs), which need to be more accurately geolocated when viewed from nearby; and a need for smaller survey areas for calibration, because it is more difficult to obtain data over large areas with mobile ground vehicles. This paper addresses these requirements by

providing a novel method to estimate line scanning camera pose with respect to the platform body frame, where the location and orientation of the platform is itself provided in world coordinates from a navigation system. The method uses the data from the navigation and line scanning camera only, avoiding the need for auxiliary sensors.

Extrinsic calibration for 2D frame cameras has been studied extensively due to their ubiquitous use across many different fields, and established solutions exist [11–14]. Calibration of 1D cameras has not received as much attention. Methods can be loosely grouped into two categories: scan-based calibration and line-based calibration [6]. Scan-based calibration requires an accurate rig with a linear actuator that moves the camera orthogonally to the line scan at a constant speed over a calibration pattern, such as a checker board [15,16]. This method is suitable for industrial inspection applications in a controlled laboratory or factory setting, where a linear actuator, manipulator arm or other rig is capable of moving the sensor through a precisely specified trajectory. Line-based calibration methods, on the other hand, allow calibration from a single line scan of a 3D target with a carefully designed pattern of lines [17,18]. Line-based approaches require that the dimensions of the calibration pattern are known precisely, and that the whole pattern has been imaged in one exposure. Recently, a variation of this method using multiple line scans of a planar calibration pattern has been proposed [19], and the use of an additional auxiliary frame camera has also been explored [6,20]. All the aforementioned approaches are suitable for well controlled environments: for scan-based calibration the movement of the sensor needs to be accurately controlled, while for line-based methods, the position of the pattern with respect to the sensor is critical. However, in a mobile ground based field platform, where the camera is rigidly mounted in a particular position to the platform, it is difficult to meet either of those requirements.

In previous methods, extrinsic parameters are usually determined with respect to the calibration pattern or an auxiliary frame camera. Therefore to determine the camera to navigation system transform either requires accurate knowledge of a pattern or points in world coordinates or an additional step such as “hand-eye” calibration [21]. Hand-eye calibration involves determining the transformation from a camera to an end effector (a robotic hand for instance), where these are rigidly linked, and is a thoroughly covered topic in the robotics literature. The problem is generally solved by imaging a calibration pattern from many different locations, where the transformations between the different end effector positions and camera to calibration pattern transformations are known using standard frame camera calibration techniques. Comparisons can be made with the problem in this paper, where the navigation system positions (and therefore any transformations between them) are known, and camera to calibration pattern transformations can be determined using any of the previously discussed methods.

As remote sensing most commonly involves imaging from an aerial or satellite platform, translation (lever arm) offsets have a smaller effect on imaging accuracy, and can be measured manually [22,23]. Accurately geolocated GCPs are commonly used to determine boresight alignment [24], which can also be adopted for ground based applications [25]. Efforts have been made to avoid the use of GCPs, by detecting points of interest in separate scans of the same area and determining their 3D position using a known digital elevation model (DEM) [22]. Similarly, non-surveyed tie-points between overlapping acquisition runs have been used in combination with bundle adjustment to determine boresight parameters [26]. The use of GCPs has also been combined with DEMs to improve accuracy and allow self-calibration [27]. Frame cameras have been used to aid in determination of boresight misalignments [28], and additionally in combination with a DEM [29]. Frame camera images have also been used to improve the geometric characteristics of processed hyperspectral linescan images from a UAV [30].

This paper provides a method for the determination of the relative 6 DOF pose of a rigidly mounted line scanning camera with respect to a navigation system on a ground based mobile platform. With this approach many of the previously outlined requirements and limitations are mitigated:

- The dimensions of the calibration pattern do not need to be known, and so it does not need to be printed to any particular accuracy, nor even measured.

- GCPs do not need to be surveyed.
- Auxiliary sensors, such as 2D frame cameras, are not required to aid the calibration.
- A single, compact calibration pattern can be used rather than widely distributed GCPs.
- Translational (lever arm) offsets are determined in addition to rotations (boresight), due to their increased significance when at close proximity to the scene.

The remainder of the paper is organised as follows. In Section 2 the theory of the proposed method is outlined in detail. Then Section 3 provides practical implementation details and the experimental method. Experimental results using the Ladybird and Shrimp robotic platforms are produced and discussed in Sections 4 and 5.

## 2. Overview of Approach

In this section, the theoretical approach used for estimating the camera pose with respect to the platform body is outlined in detail. Initially, an overview of the line scanning camera model is provided, which is an adaptation of the widely used pinhole model. This allows defining lines or rays in 3D space that intersect both the camera centre and a pixel on the sensor. When combined with the Cartesian coordinate transformations between camera, body and world frames, rays can be projected onto a surface, and conversely a world 3D point can be reprojected to a point on the 2D sensor. It is desirable to minimise any errors in the camera pose, as they directly affect mapping accuracy.

We propose a method that estimates the relative camera pose using image and navigation system data. The data are obtained by moving the platform in order to observe a calibration pattern with multiple point targets from different perspectives. The calibration pattern point locations are then manually labelled in the image data. Starting from an initial hand measured camera pose, image pixel locations of the observed pattern points and corresponding platform poses are combined, and all of the resulting rays are used to triangulate the pattern point locations in world coordinates. These point estimates are then reprojected to the sensor frame for each observation. The reprojection error is calculated as the distance between each observed and reprojected pixel. The reprojection error uncertainty is calculated by propagating the input uncertainties through each calculation as variance-covariance matrices (henceforth referred to as covariance matrices for brevity). Assuming a normal distribution of the reprojection error over input parameters, the likelihood of the data given a relative camera pose hypothesis can be estimated. By maximising the likelihood, the six relative camera pose parameters can be optimised. Following this, a random sampling based procedure is provided to estimate the uncertainties of the optimal camera pose using Markov Chain Monte Carlo (MCMC).

Throughout this paper, superscripts represent the reference frame of a particular variable. Subscripts refer to a descriptor (e.g., which pose is being referred to), axis reference, and instance identifiers for that variable, in that order. For example,  $r_{c,x}^b$  refers to the camera centre location along the  $x$  axis relative to the body frame.

### 2.1. Line Scanning Camera Model

Using the pinhole camera model with homogeneous coordinates, a point  $\mathbf{p}^w = [x, y, z, 1]^T$  in world coordinates is projected to the camera sensor at  $[u, v, 1]^T$  with the following Equation [31]:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} s = \mathbf{P} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \quad (1)$$

where  $s$  is a scale factor and  $\mathbf{P}$  can be broken down into,

$$\mathbf{P} = \mathbf{K} \mathbf{R}_c^{w-1} [\mathbf{I}_{3 \times 3} | -\mathbf{p}_c^w]. \quad (2)$$

$\mathbf{R}_c^w$  is the rotation matrix of the camera with respect to the world frame. Joined horizontally are  $\mathbf{I}_{3 \times 3}$  and  $\mathbf{p}_c^w$ , which are the identity matrix and the world camera position (i.e., the camera centre  $[r_{c,x}^w, r_{c,y}^w, r_{c,z}^w]^T$ ) respectively.  $\mathbf{K}$  is the intrinsic camera matrix:

$$\mathbf{K} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

where  $f$ ,  $u_0$  and  $v_0$  are the focal length (in pixels) and principal points respectively (we neglect skew because there is only one spatial axis). For a line scanning camera, we assume that  $v_0 = 0$  and so it follows that for a 3D world point to be visible in the 1D pixel array, it must be located near the plane that intersects the scan line on the sensor (i.e., where  $v = 0$ ) and the camera centre (focal point). How closely a point must be located to that plane depends on the instantaneous field of view (IFOV) and distance from the sensor. The IFOV is the angle over which each pixel is sensitive to radiation. While linescan image data is by definition at  $v = 0$ , reprojection errors can occur in both  $u$  and  $v$  as will be shown later. Therefore, even though the model allows for two spatial dimensions on the image sensor, it describes the projection of points for individual 1D line scan frames only.

Each pixel point  $[u, v, 1]^T$  maps to a ray or line in 3D space, which connects the sensor pixel, camera centre and object being viewed. While that ray may be defined by any two points that lie on it, the following are mathematically convenient to obtain: the camera centre  $\mathbf{p}_c^w$  and  $\mathbf{p}_s^w = \mathbf{P}^+[u, v, 1]^T$ , where  $\mathbf{P}^+$  is the pseudo-inverse of  $\mathbf{P}$  [31].

## 2.2. Rotation and Transform Conventions

In this paper, we use both Euler and axis-angle conventions to represent rotations compactly. The navigation system on the platforms used in this work provide platform pose estimates using the Euler  $zyx$  intrinsic convention (also known as Tait-Bryan or yaw-pitch-roll), which are represented as  $[\phi_x, \phi_y, \phi_z]$ , and may be converted to rotation matrices as per Berner [32] or Section 3.1 in Underwood [33]. While Euler angle representations are commonly used in robotics applications, they present the following ambiguities. Some different combinations of  $\phi_x$ ,  $\phi_y$  and  $\phi_z$  can represent the same rotation [34]. Similarly, a small freedom of rotation about a non-orthogonal axis can result in a large correlated degree of freedom spread over two Euler angles, which is difficult to interpret when estimating parameter uncertainty. For these reasons, while navigation and hand measured pose data is provided as Euler angles, we favour the axis-angle representation for all internal calculations and results. An axis-angle rotation is given as a unit length vector  $\mathbf{e}$  and a rotation  $\theta$  around it:

$$(\theta, \mathbf{e}) = \left( \theta, \begin{bmatrix} e_x \\ e_y \\ e_z \end{bmatrix} \right). \quad (4)$$

Since rotations only have three degrees of freedom, an axis-angle rotation may be expressed as a length three vector:

$$\theta \mathbf{e} = \begin{bmatrix} \theta e_x \\ \theta e_y \\ \theta e_z \end{bmatrix}. \quad (5)$$

Axis-angle rotations may be converted to rotation matrices as follows [32]:

$$\mathbf{R} = \left[ \mathbf{I}_{3 \times 3} + \sin(\theta) \mathbf{S}_n + (1 - \cos(\theta)) \mathbf{S}_n^2 \right], \quad (6)$$

where



$$\mathbf{S}_n = \begin{bmatrix} 0 & -e_z & e_y \\ e_z & 0 & -e_x \\ -e_y & e_x & 0 \end{bmatrix}. \quad (7)$$

A complete 6 DOF pose transform can be compactly represented with the three translation and three orientation parameters:

$$\mathbf{t} = [r_x, r_y, r_z, \phi_x, \phi_y, \phi_z]^T, \quad (8)$$

or

$$\mathbf{t} = [r_x, r_y, r_z, \theta e_x, \theta e_y, \theta e_z]^T, \quad (9)$$

depending on whether Euler or axis-angle conventions are used. The pose transforms of importance in this paper are the world to platform body transform  $\mathbf{t}_b^w$ , platform body to camera transform  $\mathbf{t}_c^b$ , and, combining these, the world to camera transform  $\mathbf{t}_c^w$  (see Figure 1). Note the sub- and superscripts: e.g.,  $\mathbf{t}_c^b$  denotes the translation and rotation of the camera axes with respect to the platform body.

By splitting the world pose of the camera  $\mathbf{t}_c^w$  into a combination of the body pose  $\mathbf{t}_b^w$  and the camera relative pose  $\mathbf{t}_c^b$ ,  $\mathbf{P}$  from Equation (2) can be shown as a function of the camera rotation and translation with respect to the body frame  $\mathbf{R}_c^b$  and  $\mathbf{p}_c^b$ , and platform body rotation and translation with respect to the world frame  $\mathbf{R}_b^w$  and  $\mathbf{p}_b^w$ :

$$\mathbf{P} = \mathbf{K} \left( \mathbf{R}_c^{b-1} [\mathbf{I}_{3 \times 3} | -\mathbf{p}_c^b] \right) \left( \left[ \begin{array}{c|c} \mathbf{R}_b^{w-1} & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array} \right] \left[ \begin{array}{c|c} \mathbf{I}_{3 \times 3} & -\mathbf{p}_b^w \\ \hline \mathbf{0} & 1 \end{array} \right] \right) \quad (10)$$

In our case,  $\mathbf{R}_b^w$  and  $\mathbf{p}_b^w$  are provided by the navigation system, and  $\mathbf{R}_c^b$  and  $\mathbf{p}_c^b$  are the relative camera pose parameters we would like to estimate.

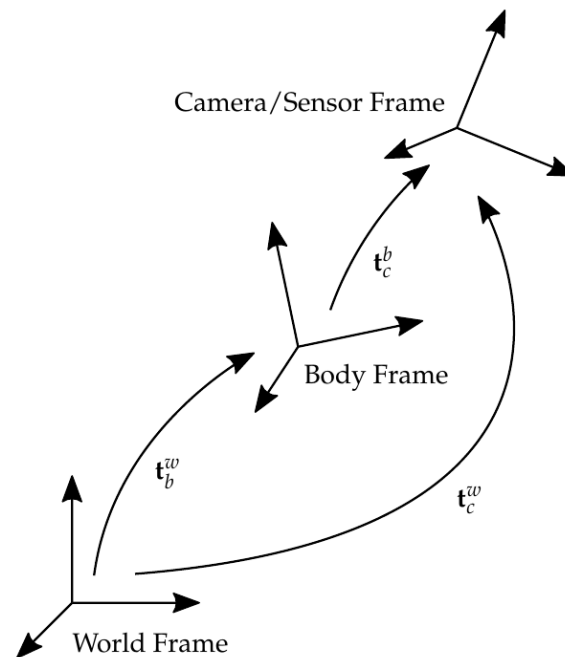
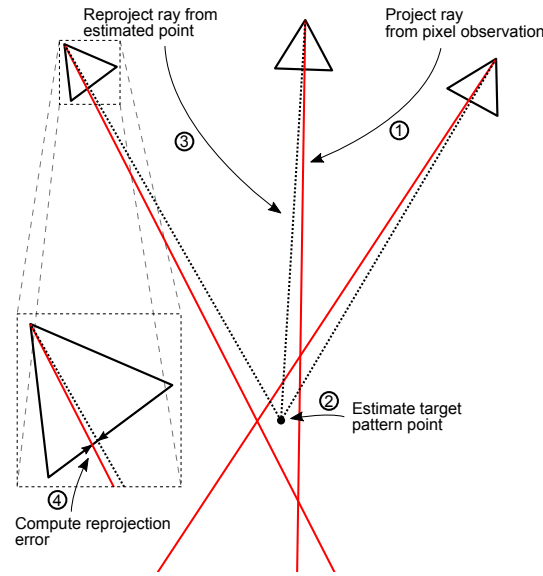


Figure 1. Summary of transforms referenced in this paper.

### 2.3. Estimation of Calibration Pattern Points

The first step of the proposed method involves estimating the location of calibration pattern points in world coordinates, as these are unknown and must be computed from the data. As shown in Figure 2, rays are calculated for each pixel observation of each calibration pattern point, given the concurrent navigation system solution and camera pose proposal. Average point locations are determined by triangulating all rays corresponding to the same calibration pattern point. Uncertainties for all inputs (pixel locations, navigation solutions and intrinsics) in the form of covariance matrices are propagated using the Jacobian of the point calculation function, yielding an uncertainty estimate (covariance matrix) for each calibration point estimate.



**Figure 2.** Method summary. Rays corresponding to individual calibration pattern point observations are determined from pixel observations and camera poses. Calibration pattern point locations are then triangulated from all rays, and subsequently reprojected to the camera sensor. A reprojection error can then be computed by calculating the difference between the reprojected point and the pixel observation. Uncertainties are propagated through at each step, which facilitates the calculation of the uncertainty for the reprojection error, and subsequently a likelihood value, which is maximised by the optimiser.

The proposed method starts with repeated imaging of points that can be uniquely identified. The use of a regular calibration pattern ensures points can be easily distinguished and is therefore recommended. The location of each pattern point  $\mathbf{p}_k^w$  for  $k \in \{1, 2, \dots, M\}$  is estimated from all of its observation rays  $i \in \{1, 2, \dots, N\}$ . There are  $M$  points on the calibration pattern and the whole pattern is viewed  $N$  times. For each point, we calculate the nearest points between all pairs of observation rays  $(i, j)$  and apply a weighted average. Nearest points between rays are calculated as follows [35]:

$$\mathbf{p}_{k,ij}^w = \mathbf{p}_{c,k,i}^w + \frac{(\mathbf{p}_{c,k,j}^w - \mathbf{p}_{c,k,i}^w) \cdot \mathbf{n}_{k,ij}}{(\mathbf{p}_{s,k,i}^w - \mathbf{p}_{c,k,i}^w) \cdot \mathbf{n}_{k,ij}} (\mathbf{p}_{s,k,i}^w - \mathbf{p}_{c,k,i}^w), \quad (11)$$

where

$$\mathbf{n}_{k,ij} = (\mathbf{p}_{s,k,j}^w - \mathbf{p}_{c,k,j}^w) \times [(\mathbf{p}_{s,k,i}^w - \mathbf{p}_{c,k,i}^w) \times (\mathbf{p}_{s,k,j}^w - \mathbf{p}_{c,k,j}^w)]. \quad (12)$$

We could estimate  $\mathbf{p}_k^w$  as the unweighted mean of all  $\mathbf{p}_{k,ij}^w$  for a given pattern point  $k$ , but some estimates are more certain than others given the conditions of how they were measured. A more accurate estimate is obtained using a weighted average according to the uncertainty. The uncertainty of each point  $\mathbf{p}_{k,ij}^w$  can be obtained by computing its Jacobian  $\mathbf{J}_{\mathbf{p}_{k,ij}^w}$  with respect to all input values. Also required are the uncertainties of the pixel and platform pose observations for each ray, expressed as covariance matrices,  $\mathbf{Q}_{uv,k,i}$ ,  $\mathbf{Q}_{\mathbf{t}_{b,k,i}^w}$ ,  $\mathbf{Q}_{uv,k,j}$  and  $\mathbf{Q}_{\mathbf{t}_{b,k,j}^w}$ , as well as intrinsic and extrinsic parameter covariances,  $\mathbf{Q}_{int}$  and  $\mathbf{Q}_{\mathbf{t}_c^b}$ . Although line scan cameras have only one pixel coordinate ( $u$ ), there is also uncertainty in the second coordinate  $v$ , because a point elicits a pixel response if it is located within the camera's IFOV, not necessarily directly on the scan line.  $\mathbf{Q}_{int}$  and  $\mathbf{Q}_{\mathbf{t}_c^b}$  contain variances and covariances of the intrinsic camera parameters and the relative camera pose respectively. All the input covariance matrices are combined into one matrix  $\mathbf{Q}_{k,ij}$ :

$$\mathbf{Q}_{k,ij} = \begin{bmatrix} \mathbf{Q}_{uv,k,i} & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{Q}_{\mathbf{t}_{b,k,i}^w} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{Q}_{uv,k,j} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathbf{Q}_{\mathbf{t}_{b,k,j}^w} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{Q}_{int} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{Q}_{\mathbf{t}_c^b} \end{bmatrix}. \quad (13)$$

No correlation between the navigation solutions of the two rays is assumed, which is reasonable if the two observations are sufficiently separated in time.  $\mathbf{Q}_{k,ij}$  and  $\mathbf{J}_{\mathbf{p}_{k,ij}^w}$  can now be used to compute the uncertainty of  $\mathbf{p}_{k,ij}^w$  (Equation (11)) as covariance matrix  $\Sigma_{\mathbf{p}_{k,ij}^w}$ :

$$\Sigma_{\mathbf{p}_{k,ij}^w} = \mathbf{J}_{\mathbf{p}_{k,ij}^w} \mathbf{Q}_{k,ij} \mathbf{J}_{\mathbf{p}_{k,ij}^w}^T. \quad (14)$$

Because we wish to estimate both  $\mathbf{t}_c^b$  and  $\mathbf{Q}_{\mathbf{t}_c^b}$  with respect to all error sources other than the camera pose, we set all elements of the  $6 \times 6$  covariance matrix  $\mathbf{Q}_{\mathbf{t}_c^b}$  to zero temporarily [33,36]. Each point  $\mathbf{p}_k^w$  on the calibration pattern can then be estimated by computing an average that is weighted according to the covariances [37]:

$$\mathbf{W}_{k,ij} = \Sigma_{\mathbf{p}_{k,ij}^w}^{-1}, \quad (15)$$

$$\Sigma_{\hat{\mathbf{p}}_k^w} = \left( \sum_i^N \sum_j^N \mathbf{W}_{k,ij} \right)^{-1}, \quad (16)$$

$$\hat{\mathbf{p}}_k^w = \Sigma_{\hat{\mathbf{p}}_k^w} \left( \sum_i^N \sum_j^N \mathbf{W}_{k,ij} \mathbf{p}_{k,ij}^w \right). \quad (17)$$

This ensures that the contribution of each closest point for each ray pair ( $\mathbf{p}_{k,ij}^w$ ) is weighted according to its certainty, taking into account navigation system uncertainty or challenging viewpoint geometry (such as a small angle between the two rays).

#### 2.4. Calculation of Reprojection Error and Likelihood Optimisation

Once estimates and uncertainties of each calibration pattern point have been obtained, they are reprojected to the camera for each observation, which allows calculating an error between each of the observed pixel locations and the reprojected pixels (see Figure 2). The uncertainties of all inputs and calibration pattern point estimates are also propagated through, which yields an uncertainty value for each reprojection error. This enables the calculation of an overall likelihood value of the data given a camera pose proposal. The optimiser maximises this likelihood by varying the camera pose to arrive at an estimate.

For each observation  $i$ ,  $\hat{\mathbf{p}}_k^w$  can be reprojected according to Equation (1), given a  $\mathbf{t}_c^b$  and corresponding navigation system solution  $\mathbf{t}_{b,k,i}^w$ . The reprojection error is calculated as the Euclidean distance between the reprojected and observed pixel locations:

$$e_{k,i} = \sqrt{(u_{k,i} - \hat{u}_{k,i})^2 + (v_{k,i} - \hat{v}_{k,i})^2}. \quad (18)$$

The reprojection is two dimensional, because non-optimal  $\mathbf{t}_c^b$  can result in reprojected pixels that deviate from the one dimensional scan line ( $\hat{v}_{k,i} \neq 0$ ), but  $v_{k,i}$  is assumed to be 0. The variance of the reprojection error can also be computed using the input covariance matrix and Jacobian  $\mathbf{j}_{e_{k,i}}$ :

$$\sigma_{e_{k,i}}^2 = \mathbf{j}_{e_{k,i}}^T \mathbf{Q}_{k,i} \mathbf{j}_{e_{k,i}}, \quad (19)$$

where

$$\mathbf{Q}_{k,i} = \begin{bmatrix} \Sigma_{\hat{\mathbf{p}}_{k,i}^w} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{Q}_{uv,k,i} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{Q}_{\mathbf{t}_{b,k,i}^w} & 0 & 0 \\ 0 & 0 & 0 & \mathbf{Q}_{int} & 0 \\ 0 & 0 & 0 & 0 & \mathbf{Q}_{\mathbf{t}_c^b} \end{bmatrix}. \quad (20)$$

The Jacobian  $\mathbf{j}_{e_{k,i}}$  is lower case because it is only one dimensional in this instance, since  $e_{k,i}$  is a scalar value. As in Equation (13), we again set all elements of  $\mathbf{Q}_{\mathbf{t}_c^b}$  to zero. The log likelihood of a transform  $\mathbf{t}_c^b$  given the observations can then be estimated as,

$$\log \Lambda = - \sum_k^M \sum_i^N \frac{e_{k,i}^2}{2\sigma_{e_{k,i}}^2}. \quad (21)$$

The objective is to maximise  $\log \Lambda$ , by varying the 6-DOF  $\mathbf{t}_c^b$  vector. This can be achieved using standard optimisation methods to minimise the negative log likelihood:

$$\underset{\mathbf{t}_c^b}{\operatorname{argmin}} -\log \Lambda = \underset{\mathbf{t}_c^b}{\operatorname{argmin}} \sum_k^M \sum_i^N \frac{e_{k,i}^2}{2\sigma_{e_{k,i}}^2}, \quad (22)$$

$\log \Lambda$  is fully recalculated at each optimisation iteration, which includes the triangulation of calibration pattern points and calculation of their reprojection error.

## 2.5. Variance-Covariance Matrix Estimation

Once the relative camera pose  $\mathbf{t}_c^b$  has been determined, it is desirable to approximate the covariance matrix of the solution, which provides an estimate of how uncertain the six relative camera pose parameters are. In combination with covariances of other parameters, such as the navigation system solution, this also allows mapping accuracy to be quantified. In other words, the result provides values for  $\mathbf{Q}_{\mathbf{t}_c^b}$ , completing the full covariance matrix (see Equation (29) in Section 3.6). Note that all elements of  $\mathbf{Q}_{\mathbf{t}_c^b}$  are set to zero for its estimation and optimisation, as previously mentioned in Sections 2.3 and 2.4.

The proposed approach is based on similar work done with lidar sensors [33,36], but the details differ because 1D cameras do not directly provide depth information. We propose a random sampling based method, where a set of sample sensor to body transforms are selected using a MCMC algorithm [38], which differs from the Monte Carlo (MC) importance sampling approach in [33,36]. This provides greater sampling efficiency and avoids the need to manually define a sampling region. The algorithm is guided by the likelihood of each relative camera pose sample, which governs the selection of the next sample.

There are several MCMC variations, but they all share the property that each sample is selected based on the previous. For a large number of samples, the distribution tends towards the probability

distribution that is being sampled from (i.e.,  $\Lambda$  in this paper) [39]. For further details about MCMC sampling, the reader is referred to the numerous resources available on the topic [38,39]. The MCMC algorithm provides a list of samples  $\{t_{c,1}^b, t_{c,2}^b, \dots, t_{c,l}^b, \dots, t_{c,r}^b\}$ , which are distributed according to  $\Lambda$ , from which the covariance can be computed as:

$$\mathbf{Q}_{t_c^{b*}} = \frac{1}{r-1} \sum_{l=1}^r \left( \mathbf{t}_{c,l}^b - \bar{\mathbf{t}}_c^b \right) \left( \mathbf{t}_{c,l}^b - \bar{\mathbf{t}}_c^b \right)^T, \quad (23)$$

where

$$\bar{\mathbf{t}}_c^b = \frac{1}{r} \sum_{l=1}^r \mathbf{t}_{c,l}^b, \quad (24)$$

### 3. Materials and Methods

This section outlines the equipment and methods used to obtain the data and analyse the results. A planar calibration pattern was placed in the environment and imaged from several different orientations using a line scanning camera mounted to two different ground based robotic platforms. A navigation system mounted to each platform recorded the 6 DOF position and orientation of the platforms ( $\mathbf{t}_b^w = [r_{b,x}^w, r_{b,y}^w, r_{b,z}^w, \phi_{b,x}^w, \phi_{b,y}^w, \phi_{b,z}^w]$ ) throughout the acquisition period. Image pixel locations of calibration pattern points and matching robot poses were then used to estimate the relative camera pose using an iterative optimisation algorithm. Finally, the uncertainty of the camera pose estimate in the form of a covariance matrix was approximated using MCMC.

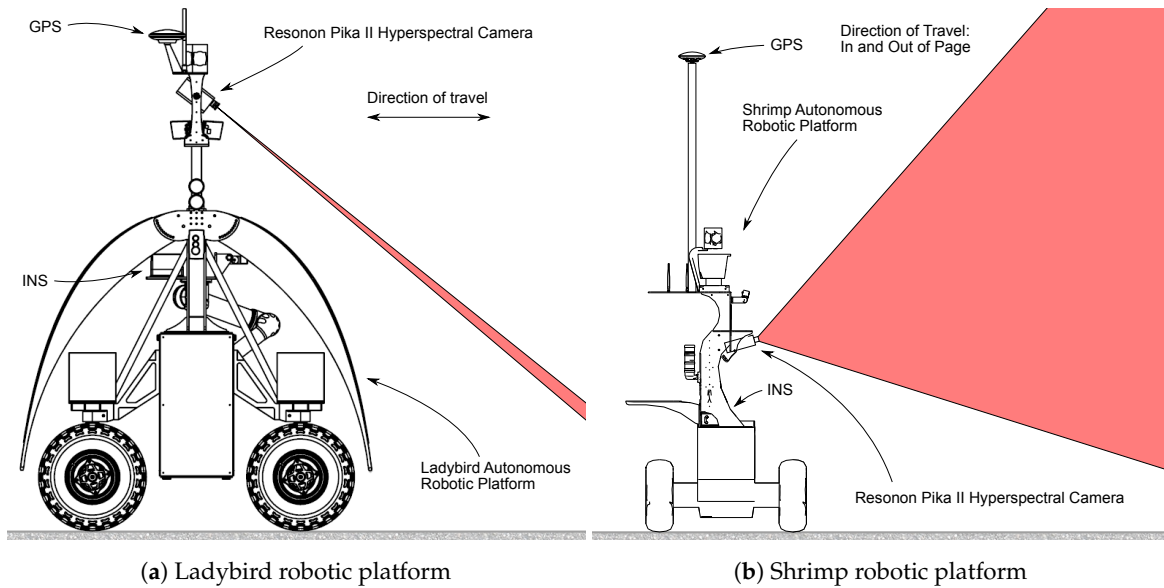
First the ground based mobile platforms and associated sensors used to acquire data are introduced, followed by a description of the data acquisition process and extraction of pattern point observations. The implementation of the method presented in Section 2 is outlined, which includes the optimisation and an outlier removal process. Methods for mapping image data and comparing camera poses are presented, as required for the analysis of the results, and a method is presented to calculate the basin of attraction, to assess the sensitivity of the process to the initial camera pose.

#### 3.1. Mobile Sensing Platforms

A line scanning hyperspectral camera was mounted to two different robotic platforms, Ladybird and Shrimp (Figure 3). Both were designed and built at the Australian Centre for Field Robotics (ACFR) at The University of Sydney as flexible tools to support a range of research applications [40–45]. The sensor suite on both platforms includes a real time kinematic (RTK)/global positioning system (GPS)/inertial navigation system (INS), which provides platform pose and covariance estimates (details in Table 1). The GPS units on both platforms are identical, but the Shrimp platform uses a lower grade inertial measurement unit (IMU) than the Ladybird platform.

Line scan image data were acquired with a Resonon Pika II visible to near infrared (VNIR) line scanning camera that was mounted to the Ladybird and Shrimp robots in a push broom configuration. For the Ladybird, the camera was oriented such that the scan line is horizontal, pitched down for scanning the ground surface (Figure 3a). On Shrimp, the camera was mounted such that the scan line is vertical, and pitched upwards slightly to allow scanning of upright objects (Figure 3b). The camera produces hyperspectral images of 648 spatial by 244 spectral pixels (spectral resolution of 2 nm from 390.9–887.4 nm) at a rate of 133 frames per second and native bit depth of 12. For the purposes of this paper, the spectral dimension was averaged to produce 648 pixel monochrome scan lines. Apart from this averaging step, the method described in this paper is not particular to hyperspectral cameras and may be applied equally to other types of line scanning imagers. Schneider Cinegon 6 mm and 8 mm objective lenses were used for Shrimp and Ladybird respectively, and manually focused with a checker board at the typical distance to the scene. The principal point of the camera/lens combination was assumed to be at the centre of the line scan ( $u_0 = 323$ ), the focal length was assumed to be as

per the manufacturer supplied measurements (see Table 1), and distortion was assumed to be zero. Hand measured pose estimates and manufacturer supplied lens details are shown in Table 1.



**Figure 3.** The Ladybird (a) and Shrimp (b) robotic platforms and sensor configurations.

**Table 1.** Platform configurations.

	Ladybird	Shrimp
<b>Manually Measured Camera Pose <math>t_c^b</math></b>		
$r_{c,x}^b, r_{c,y}^b, r_{c,z}^b$ (m)	0.2, 0.0, −0.8	0.0, −0.2, −0.5
$\phi_{c,x}^b, \phi_{c,y}^b, \phi_{c,z}^b$ (°)	−56.0, 0.0, −90.0	0.0, 105.0, −90.0
<b>Camera Lens Details</b>		
Manufacturer	Schneider	Schneider
Model	Cinegon 8 mm	Cinegon 6 mm
Focal length	8.2 mm	6.2 mm
Approx. aperture	f/2.5	f/3.0
IFOV	1.88 mrad	2.5 mrad
<b>Navigation System Details</b>		
Manufacturer	Novatel	Novatel
GPS receiver	ProPak-G2plus	ProPak-G2plus
IMU	Honeywell HG1700	IMU-CPT

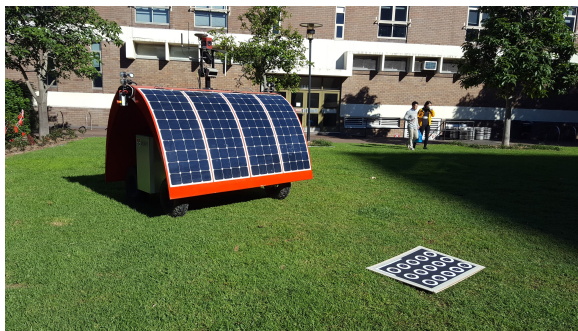
Initial pose estimates were measured by hand with the mobile platforms on a level surface using measuring tape for translational offsets, and a digital inclinometer (SPI Pro 3600) for angular offsets around the robots' horizontal  $x$  and  $y$  axes. Angular offsets around the robots' vertical  $z$  axis were assumed to be the intended mounting orientations, which are in increments of  $90^\circ$  for both platforms. Note that if the camera is mounted at angles that are clearly not in  $90^\circ$  increments, referring to a CAD model is recommended. Hand measured translation parameters ( $r_{c,x}^b$ ,  $r_{c,y}^b$  and  $r_{c,z}^b$ ) were assumed to have a standard deviation ( $\sigma$ ) of 0.1 m and orientation parameters ( $\phi_{c,x}^b$ ,  $\phi_{c,y}^b$  and  $\phi_{c,z}^b$ ) were assumed to have a  $\sigma$  of  $2^\circ$ .

### 3.2. Data Acquisition

A calibration pattern with 15 points arranged in a  $3 \times 5$  pattern was printed to an A1 size sheet of paper and mounted to a flat rigid plywood board (see Figure 4c). The pattern was designed to maximise contrast for efficient extraction of pattern points. A corner shape was added to one side of



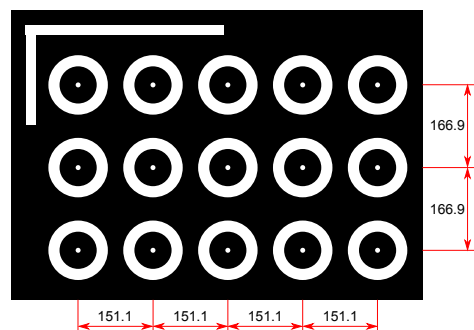
the pattern to facilitate unique identification of each point. It is not necessary to know the pattern's dimensions for recovery of the platform to camera pose, as each point is treated independently during the calibration. This also means that theoretically a single point with sufficient observations could be used for calibration. However, we added more pattern points since there is no significant practical cost, efficiently increasing the amount of data obtained. For the Ladybird platform, the pattern board was placed on relatively flat ground (see Figure 4a). As shown in Figure 5, the pattern was scanned from several directions around a circle with the calibration pattern in the centre. Two types of scans were performed, one with the robot's wheels flat on the ground and one with one side of the robot elevated by driving over an aluminium channel. This raised two of the wheels by approximately 100 mm, inducing a roll of approximately  $4^\circ$ . For the Shrimp platform, the same calibration pattern was mounted to a ladder in an approximately vertical orientation (see Figure 4b). In this case data were acquired next to a hill with various orientations and positions with respect the pattern, where the hill caused continuously variable roll and pitch, up to approx.  $17^\circ$  (see Figure 5b). For both platforms, body orientation was intentionally varied as much as possible in an attempt to maximise observability of parameters [33]. The robots were manually operated throughout the acquisition period, and care was taken to move slowly and smoothly while the calibration pattern was imaged.



(a) Ladybird data acquisition



(b) Shrimp data acquisition



(c) Calibration pattern

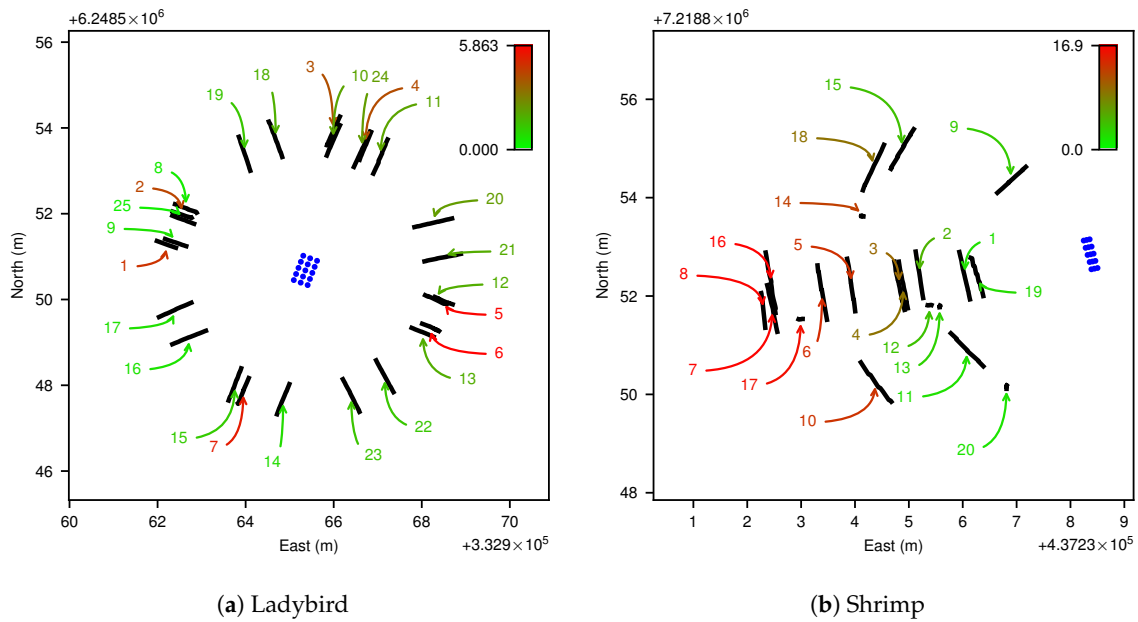
**Figure 4.** Data acquisition configuration and location for Ladybird (a) and Shrimp (b). The calibration pattern used in both instances is shown in (c). Dimensions are in mm. The centre dots are 10 mm in diameter, sized to be as small as possible while still being visible in the image data in order to maximise labelling accuracy. The outer rings help with locating the points in the data for labelling, and are 75–120 mm, inner to outer diameter. Note that these dimensions do not need to be known for the optimisation procedure, nor is it necessary to print them in any particular strict arrangement, though a regular pattern is recommended to allow easy identification of each point.

All data was timestamped allowing association between individual scan lines and platform pose solutions. Localisation uncertainties reported by the navigation system are shown in Table 2 as median standard deviations (i.e., square root of the diagonals of the covariance matrices only) for

the acquisition runs, which illustrates that the navigation system in the Ladybird platform is able to provide body pose estimates with much greater certainty than the navigation system on Shrimp, due to the higher grade IMU.

**Table 2.** Median navigation system uncertainties as 1 standard deviation.

Platform	$\sigma_{r_{b,x}^w}$ (m)	$\sigma_{r_{b,y}^w}$ (m)	$\sigma_{r_{b,z}^w}$ (m)	$\sigma_{\phi_{b,x}^w}$ (°)	$\sigma_{\phi_{b,y}^w}$ (°)	$\sigma_{\phi_{b,z}^w}$ (°)
Ladybird	$1.052 \times 10^{-2}$	$1.305 \times 10^{-2}$	$1.118 \times 10^{-2}$	$2.362 \times 10^{-1}$	$2.636 \times 10^{-1}$	$1.053 \times 10^{-1}$
Shrimp	$4.520 \times 10^{-2}$	$4.369 \times 10^{-2}$	$4.887 \times 10^{-2}$	$7.534 \times 10^{-1}$	$7.284 \times 10^{-1}$	$8.416 \times 10^{-1}$



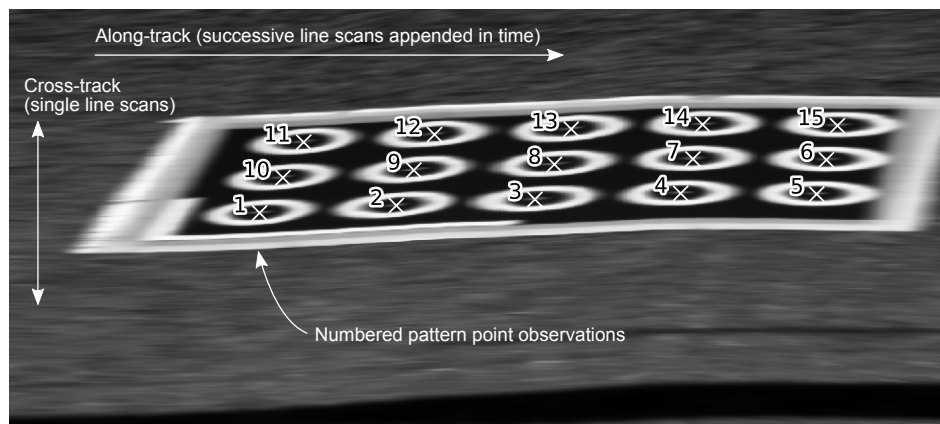
**Figure 5.** Top down view of platform positions (from navigation system) during each observation of the calibration pattern for Ladybird (a) and Shrimp (b). Each observation provides a complete view of all points on the calibration pattern, as well as concurrent navigation system solutions. The calibration pattern points are indicated with blue dots. The observation runs are numbered, and coloured according to the norm of vehicle body pitch and roll (i.e.,  $\sqrt{\phi_{b,x}^w{}^2 + \phi_{b,y}^w{}^2}$ ), as indicated by the colour bars in degrees. Note that in (b) the calibration pattern was upright (mounted to a ladder), which is why the points appear more closely clustered from the top-down perspective. Some of the observation runs appear very short in (b). This is because in these instances the platform scanned the pattern by rotating on the spot.

### 3.3. Pattern Pixel Extraction

Approximate pixel locations of points on the calibration pattern were selected manually by appending successive line scans to form a rectangular image and selecting individual pattern points in order (see Figures 6 and 7). Note that line scans were concatenated naively, ignoring camera or body pose data (i.e., not mapped or georeferenced). This worked well for this purpose, because the platforms were moved slowly and smoothly while the calibration pattern was scanned. Particular care was taken to ensure that point ID numbers were consistent for all observations of the calibration pattern. Pixel locations were then refined to sub pixel precision by extracting a  $10 \times 10$  patch around the selected points and resizing it to  $100 \times 100$  pixels using bi-cubic interpolation. The intensity peak closest to the centre was taken as the pattern point pixel location. Along-track, the closest time stamp was used to obtain the corresponding navigation solution. This provides pixel position  $u_{k,i}$  and platform pose  $[r_{b,x}^w, r_{b,y}^w, r_{b,z}^w, \phi_{b,x}^w, \phi_{b,y}^w, \phi_{b,z}^w]^T$ , which are necessary for calibration according to Equation (22).



**Figure 6.** Example grey scale image obtained by appending successive line scans from Ladybird (without pose compensation). A single view of all points on the calibration pattern is referred to as an observation in this paper, including concurrent navigation system data. The red box indicates one such observation, shown as a more detailed close-up view in Figure 7. The two views of the calibration pattern shown in this figure were obtained by driving the platform forward, producing the first view, and then backwards, giving the second view, which is therefore a mirrored version of the first. The second view appears more stretched because it was scanned more slowly, generating more line scans, given the same fixed frame rate. Note that the observations shown in Figure 5a only include one observation from each forward-reverse pair, because they provide almost the same information (i.e., similar navigation system solutions).



**Figure 7.** Close-up view of the observation indicated in Figure 6, showing manually labelled calibration point locations (white crosses) with ID numbers. Each 648 pixel line scan is vertical and concatenated horizontally. A similar image was generated for each observation of the calibration pattern prior to manually labelling each point. Care was taken to ensure the numbering scheme remained consistent for all observations of the pattern.

### 3.4. Optimisation and Uncertainty Estimation

Optimisation was performed using the Powell optimiser algorithm provided by the SciPy python package [46,47]. While other optimisers may be suitable for this task, as long as they minimize a scalar (negative log likelihood), while varying a vector (relative camera pose), the Powell algorithm achieved acceptable performance with the following tolerance values:  $tolx = 1 \times 10^{-5}$  and  $ftol = 1 \times 10^{-8}$ . The objective function that was provided to the optimiser takes the relative camera pose parameters ( $\mathbf{t}_c^b = [r_{c,x}^b, r_{c,y}^b, r_{c,z}^b, \theta_{c,x}^b, \theta_{c,y}^b, \theta_{c,z}^b]^T$ ) and computes the negative log likelihood  $-\log\Lambda$  (see Equation (22)) given all pixel locations and navigation system solutions. The optimiser repeatedly calls this function, updating  $\mathbf{t}_c^b$  in order to find a relative camera pose  $\mathbf{t}_c^{b*}$  that minimises  $-\log\Lambda$ .

As described in Section 2.5 we use MCMC to estimate uncertainties in the form of a covariance matrix. MCMC was performed with the emcee python package [38], which was given a function that computes the log likelihood (Equation (21)). The algorithm was initialised with the previously optimised relative camera pose  $\mathbf{t}_c^{b*}$ , and run with 250 walkers and 100 iterations, yielding 25,000 samples. A burn in run was also performed with 100 iterations to allow the function to explore the local region prior to performing the actual sampling run. Each sample represents one hypothetical parameter vector  $\mathbf{t}_c^b$ . The distribution of the samples generated by the MCMC algorithm correspond to

$\log \Lambda$ , so the uncertainty of the relative camera pose estimate,  $\mathbf{Q}_{t^{b*}}$ , can be estimated by computing Equation (23).

For all calculations of  $\log \Lambda$ ,  $6 \times 6$  covariances for the platform pose were provided by the navigation system at each time stamp. It was assumed that a  $u$  pixel point location could be estimated to within one standard deviation of 0.5 pixels (i.e.,  $\sigma_u = 0.5 \text{ pixels}$ ). If a point is visible it must also be within the IFOV of the sensor (see Table 1), which is approximately 2 pixels for both platforms. We assumed this to span two standard deviations (95%), and so one standard deviation is 0.5 pixels ( $\sigma_v = 0.5 \text{ pixels}$ ). Principal point and focal length were assumed to have a standard deviation of 2 pixels and 0.1 mm respectively. As previously mentioned, the uncertainty of the relative camera pose,  $\mathbf{Q}_{t^b}$ , was temporarily set to zero (see Sections 2.3 and 2.4).

### 3.5. Outlier Removal

Unusually high reprojection errors were removed by an iterative process of outlier rejection. First optimisation was performed on all observations shown in Figure 5 for each platform. Reprojection errors,  $e_{k,i}$  were calculated for each observation  $i$  of each pattern point  $k$  (see Equation (18)). These were then averaged per observation:

$$e_i = \frac{\sum_k^M e_{k,i}}{M}. \quad (25)$$

The observation  $i$  with the largest mean reprojection error was then removed from the data set and the process was repeated several times (i.e., optimise, calculate reprojection error, remove observation with largest reprojection error). The removal process may be stopped once all mean reprojection errors are below a threshold.

### 3.6. Mapping

To demonstrate mapping performance, rays were projected to a plane that was fitted to the estimated pattern point coordinates. Utilising the method in Section 2.3, the pattern points ( $\mathbf{p}_k^w = [r_{k,x}^w, r_{k,y}^w, r_{k,z}^w]^T$ ) were first calculated given the data and relative camera pose. Using the general form of the equation of a plane  $ax + by + cz + d = 0$ , a best fit plane can be found in a linear least squares fashion (setting  $c = -1$ ):

$$\mathbf{Ax} = \mathbf{b}, \quad \begin{bmatrix} r_{1,x}^w & r_{1,y}^w & 1 \\ r_{2,x}^w & r_{2,y}^w & 1 \\ \vdots & \vdots & \vdots \\ r_{M,x}^w & r_{M,y}^w & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ d \end{bmatrix} = \begin{bmatrix} r_{1,z}^w \\ r_{2,z}^w \\ \vdots \\ r_{M,z}^w \end{bmatrix}, \quad (26)$$

where the plane parameters  $\mathbf{x}$  can be solved for by left multiplying  $\mathbf{b}$  with the pseudo-inverse of  $\mathbf{A}$ ,  $\mathbf{A}^+$ .

The rays for observation  $i$  of pattern point  $k$ , defined by  $\mathbf{p}_{c_{i,k}}^w$  and  $\mathbf{p}_{s_{i,k}}^w$  as calculated in Section 2.1, can then be projected to the plane by computing their point of intersection:

$$\mathbf{p}_{proj,i,k}^w = \frac{([0, 0, -d/c]^T - \mathbf{p}_{c_{i,k}}^w) \cdot [a, b, c]^T}{(\mathbf{p}_{s_{i,k}}^w - \mathbf{p}_{c_{i,k}}^w) \cdot [a, b, c]^T}. \quad (27)$$

Knowing all input covariance matrices, including the covariance of the relative camera position and orientation, as obtained using MCMC (see Section 2.5), the uncertainty of each point projection  $\mathbf{p}_{proj,i,k}^w$  may also be calculated. First the partial derivatives of Equation (27) with respect to all inputs were computed to yield the Jacobian  $\mathbf{J}_{p_{proj,i,k}}$  of the  $x$ ,  $y$  and  $z$  position of each point. The uncertainty of each projected point can then be calculated:

$$\Sigma_{\mathbf{p}_{proj,i,k}^w} = \mathbf{J}_{p_{proj,i,k}} \mathbf{Q}_{k,i} \mathbf{J}_{p_{proj,i,k}}^T, \quad (28)$$

where

$$\mathbf{Q}_i = \begin{bmatrix} \mathbf{Q}_{uv,k,i} & 0 & 0 & 0 \\ 0 & \mathbf{Q}_{b,k,i}^w & 0 & 0 \\ 0 & 0 & \mathbf{Q}_{int} & 0 \\ 0 & 0 & 0 & \mathbf{Q}_{t_c^b} \end{bmatrix}. \quad (29)$$

### 3.7. Comparing Poses

To assess how close a solution is to the optimal, we require a distance metric between two different 6 DOF pose transforms. As described in Section 2.2, each pose vector is composed of three translation and three orientation parameters. Given two unique poses  $\mathbf{t}_1 = [r_{1,x}, r_{1,y}, r_{1,z}, \theta_1 e_{1,x}, \theta_1 e_{1,y}, \theta_1 e_{1,z}]^T$  and  $\mathbf{t}_2 = [r_{2,x}, r_{2,y}, r_{2,z}, \theta_2 e_{2,x}, \theta_2 e_{2,y}, \theta_2 e_{2,z}]^T$ , we can compare the translation parts easily by computing their Euclidean distance:

$$d_{1,2} = \sqrt{(r_{2,x} - r_{1,x})^2 + (r_{2,y} - r_{1,y})^2 + (r_{2,z} - r_{1,z})^2}. \quad (30)$$

However, measuring the distance or difference between two rotations is more complicated, and the readers are referred to Huynh [34] for an in-depth analysis of the topic. Huynh [34] presents and recommends a number of metrics for comparing rotations. Of these, the geodesic on the unit sphere was chosen, because it represents the magnitude of the rotation angle required to align the two rotations, which was deemed to be an intuitive measure. It can be computed as follows:

$$\Phi_{1,2} = 2 \arccos(|\mathbf{q}_1 \cdot \mathbf{q}_2|), \quad (31)$$

where  $\mathbf{q}_1$  and  $\mathbf{q}_2$  are unit quaternion equivalents of  $[\theta_1 e_{1,x}, \theta_1 e_{1,y}, \theta_1 e_{1,z}]^T$  and  $[\theta_2 e_{2,x}, \theta_2 e_{2,y}, \theta_2 e_{2,z}]^T$  respectively, computed as:

$$\begin{aligned} \mathbf{q}_i &= a_i + b_i \mathbf{i} + c_i \mathbf{j} + d_i \mathbf{k} \\ &= \cos\left(\frac{\theta_i}{2}\right) + e_{i,x} \sin\left(\frac{\theta_i}{2}\right) \mathbf{i} + \\ &\quad e_{i,y} \sin\left(\frac{\theta_i}{2}\right) \mathbf{j} + e_{i,z} \sin\left(\frac{\theta_i}{2}\right) \mathbf{k}. \end{aligned} \quad (32)$$

$\Phi_{1,2}$  can also be interpreted to be equal to the absolute value of the angular magnitude  $\theta$  (in the range  $[-\pi, \pi]$ ) of the axis-angle rotation required to align the two orientations. For this reason, the metric will be simply referred to as the axis-angle difference. Combined,  $d_{1,2}$  and  $\Phi_{1,2}$  form a 2D pose distance that is convenient for visualisation.

### 3.8. Basin of Attraction

Because the method in this paper requires an approximate initial camera pose, it is important to numerically quantify how precise this initial camera pose must be to yield an accurate optimised estimate. To measure how far an initial hand measured camera pose can be from the optimum, while still resulting in correct global convergence, we test a set of starting conditions that are altered by different amounts, and measure how close the optimal result is from the known global optimum. Due to the high dimensionality of the search space, a random sub-sampling of initial poses is performed. Deviation of initial values from the known optimum is quantified in the two dimensional pose distance space defined in Section 3.7. The 2D space was sampled uniformly in a grid, and for each location a 6D initial parameter vector was randomly generated at the corresponding Euclidean and axis-angle distance from the known optimal value.



First an  $x$ ,  $y$  and  $z$  translation vector was generated at random, uniformly distributed over an equal negative to positive range for all three parameters. The vector was normalised to unity and then multiplied by the Euclidean distance value of the corresponding grid position. The resulting vector was added to the optimised translation parameters, yielding the initial coordinates. Similarly, three axis-angle orientation values were randomly generated in the same way, normalised to unity and multiplied by the corresponding axis-angle difference value at the given grid location, yielding an axis-angle rotation. The optimised orientation parameters were then rotated by this difference rotation, producing the initial orientation values.

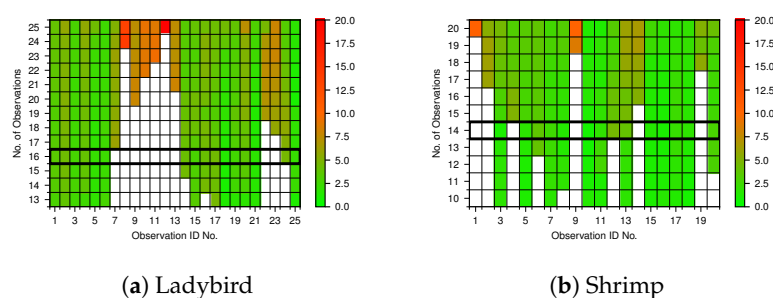
This yields a set of sparse random 6 DOF samples that are uniformly spaced in terms of pose distance from the known optimal camera pose, allowing the basin of attraction to be mapped. Optimisation was performed for each randomly generated initial camera pose on the grid. For each result, the Mahalanobis distance to the reference optimum was calculated, given the covariance matrix resulting from the MCMC uncertainty estimation on the optimised parameters.

#### 4. Results

This section presents the results of line scan camera pose estimation for two different platforms and configurations, including outlier rejection, resulting camera pose and uncertainty, in-depth analysis of the uncertainty, the impact of platform pose diversity on the accuracy, and finally the combined mapping uncertainty.

##### 4.1. Outlier Rejection

The iterative results of outlier removal based on reprojection errors are shown in Figure 8. Each row represents an outlier removal iteration, labelled by the number of remaining observations, where the top row includes all observations. Each column represents one of the observations from Figure 5, where each observation is one view of all 15 points on the calibration pattern. The colour of each cell indicates the mean reprojection error of the 15 points within the single observation of the calibration pattern, for a particular outlier removal iteration. In each figure, the black rectangle highlights the row with the greatest number of observations that all have mean errors less than a 5 pixel threshold. Ladybird exhibited a greater number of outliers and higher worst-case reprojection errors than Shrimp, with 9 compared to 6 outliers respectively. This resulted in 16 inliers for Ladybird and 14 for Shrimp.



**Figure 8.** Outlier removal and average reprojection errors for (a) Ladybird and (b) Shrimp. Each row represents an iteration as described in Section 3.5, labelled with the number of remaining observations, and each column represents an observation of the calibration pattern as per Figure 5. White cells indicate that an observation has been removed for that iteration. Otherwise, the cell colour indicates the mean reprojection error in pixels for that particular observation of the calibration pattern at a given outlier removal iteration. For example, in (a), observation No. 12 exhibited the greatest mean reprojection error with 25 observations, and was therefore the first observation to be removed. The highlighted rectangle points to the row with the greatest number of remaining observations with a mean reprojection error of less than 5 pixels.



#### 4.2. Pose and Uncertainty Results

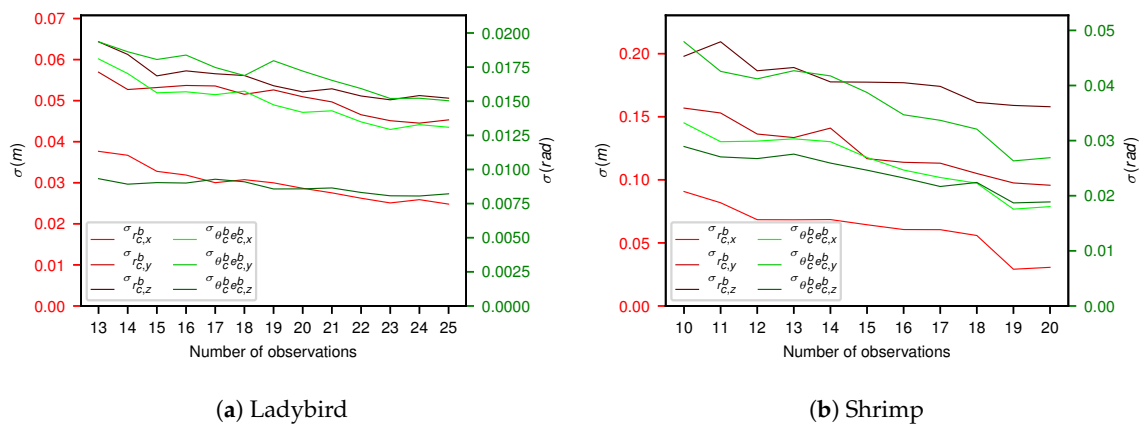
The relative camera pose transforms and associated uncertainties (one standard deviation) for both platforms are shown in Table 3 before and after outlier removal. The hand measured estimate is also shown for reference, where the tolerances reflect the difficulty of measurement. The results for the Shrimp platform exhibit greater uncertainty compared to the Ladybird platform.

**Table 3.**  $t_c^b$  estimates and uncertainties (axis-angle rotations).

	$r_{c,x}^b$ (m)	$r_{c,y}^b$ (m)	$r_{c,z}^b$ (m)	$\theta_c^b e_{c,x}^b$ (rad)	$\theta_c^b e_{c,y}^b$ (rad)	$\theta_c^b e_{c,z}^b$ (rad)
<b>Ladybird</b>						
All observations	$0.147 \pm 0.025$	$-0.128 \pm 0.045$	$-0.630 \pm 0.051$	$-0.849 \pm 0.013$	$0.768 \pm 0.015$	$-1.420 \pm 0.008$
16 observations	$0.189 \pm 0.032$	$-0.142 \pm 0.054$	$-0.794 \pm 0.057$	$-0.822 \pm 0.016$	$0.738 \pm 0.018$	$-1.429 \pm 0.009$
Hand measured	$0.200 \pm 0.100$	$0.000 \pm 0.100$	$-0.800 \pm 0.100$	$-0.762 \pm 0.039$	$0.762 \pm 0.039$	$-1.433 \pm 0.037$
<b>Shrimp</b>						
All observations	$0.044 \pm 0.031$	$-0.133 \pm 0.096$	$-0.660 \pm 0.158$	$1.409 \pm 0.018$	$1.400 \pm 0.027$	$-1.078 \pm 0.019$
14 observations	$-0.010 \pm 0.069$	$-0.080 \pm 0.141$	$-0.579 \pm 0.178$	$1.380 \pm 0.030$	$1.427 \pm 0.042$	$-1.093 \pm 0.026$
Hand measured	$0.000 \pm 0.100$	$-0.200 \pm 0.100$	$-0.500 \pm 0.100$	$1.399 \pm 0.026$	$1.399 \pm 0.088$	$-1.074 \pm 0.071$

Note: Hand measured orientation uncertainties are  $2^\circ$  for each parameter in Euler representation, converted to axis-angle representation by propagating the covariance matrix using the Jacobians of the conversion function.

In Figure 9 each outlier removal stage is plotted against each parameter's standard deviation. As would be expected, increasing the number of observations decreases the uncertainty of the estimate.

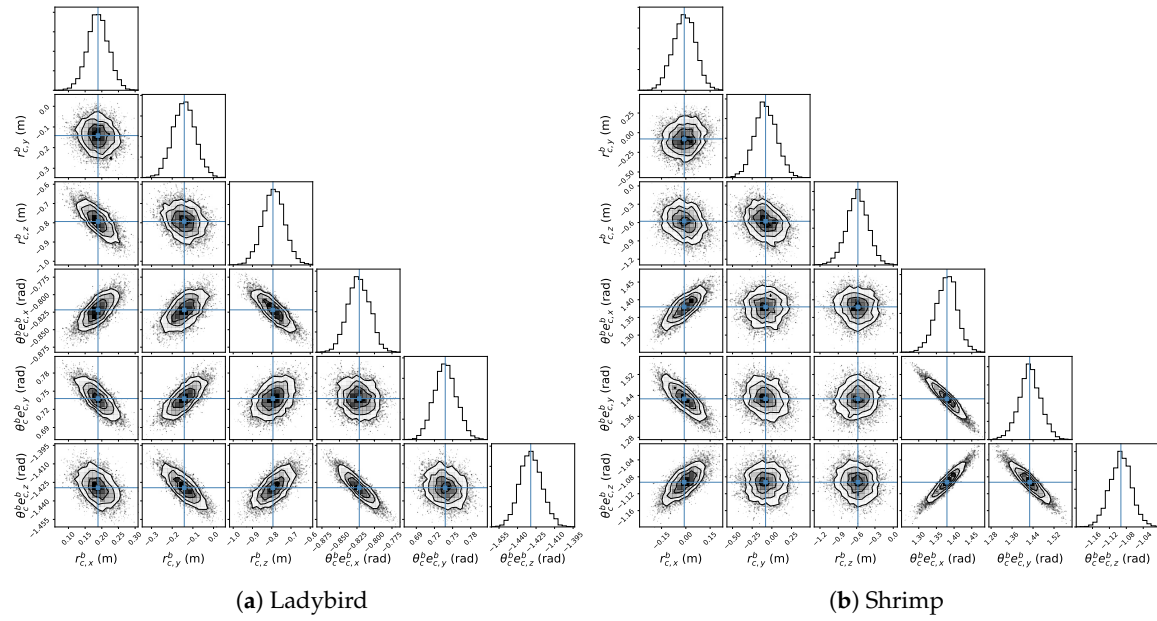


**Figure 9.** The number of observations vs. standard deviation for (a) Ladybird and (b) Shrimp respectively. The x axis mirrors the outlier removal stages shown in Figure 8.

The number of observations affects the number of computations and therefore has a significant effect on calibration and MCMC run time. For Ladybird, optimisation and MCMC took approx. 15 min and 7 h respectively for all 25 observations. For 16 observations, this was reduced to just over 7 min and 4 h. For Shrimp, the respective optimisation and MCMC times were reduced from approx. 5.5 min and 5 h for all 20 observations to just over 2 min and 3 h with 14 observations.

#### 4.3. In-Depth Uncertainty Analysis

Examining uncertainty in more detail, MCMC samples are shown on a corner plot in Figure 10 [48]. Each sub-plot below the diagonal provides a 2D histogram, showing the MCMC sample density between two parameters (i.e., the marginal likelihood distribution for a parameter pair), and on the diagonal a 1D histogram, giving the sample density for the marginal likelihood distribution for each single parameter.



**Figure 10.** Corner plots for (a) Ladybird and (b) Shrimp platforms for 16 and 14 observations respectively. Each sub-plot below the diagonal provides a 2D histogram of MCMC sample values for a pair of relative camera pose parameters. The sub-plots on the diagonal show 1D histograms for each parameter.

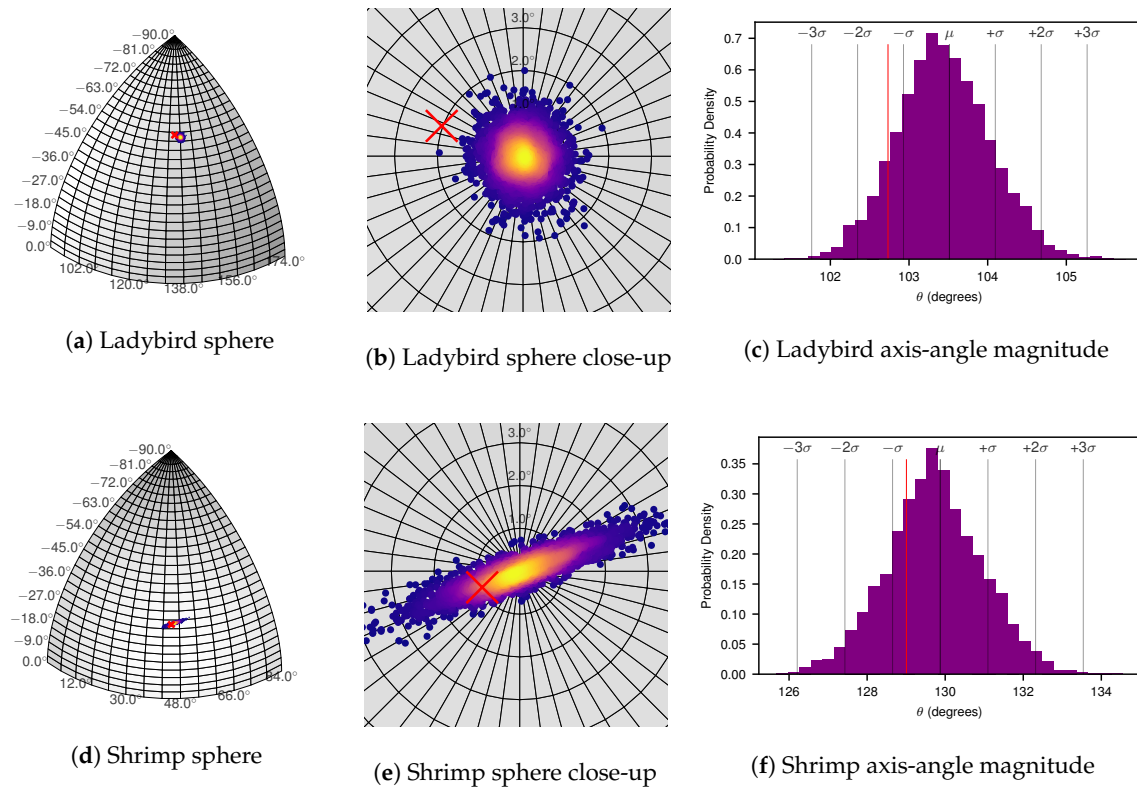
For human interpretation of the uncertainty, in Figure 11a,d the MCMC sample values for  $[\theta_{c,x}^b, \theta_{c,y}^b, \theta_{c,z}^b]^T$  are plotted on a sphere. For visualisation only, each point is coloured according to a kernel density estimate (KDE) performed on all MCMC samples to give an indication of the axis-angle vector marginal likelihood. Hand measured pose estimates are shown as red crosses. Likewise, in Figure 11c,f, values for  $\theta_c^b$  are presented in a histogram, showing the marginal likelihood of the axis-angle magnitude. While both figures indicate a clustering to within two degrees, the Shrimp platform's distribution exhibits a more elongated elliptical shape, while for Ladybird, it is more uniformly spread in all directions. Also apparent for Ladybird is that the manual measurements are well outside the region of high likelihood, both in terms of the axis-angle unit vector and magnitude. Conversely, the hand measured pose for the Shrimp platform is highly likely given the data, suggesting the initial manual measurement may have been more accurate than for Ladybird.

In Figure 12, the distributions of MCMC samples is shown superimposed on the corresponding platform model. Translation parameters are presented as a 2D histogram, similar to Figure 10, demonstrating the marginal density of the likelihood distribution from each orthogonal viewpoint. To present the orientation parameter distribution, line segments coincident with the camera's viewing direction, and anchored to the optimized camera centre, are rotated by each MCMC rotation sample. Each line is semi-transparent, and so as all samples build up, the density distribution of the camera orientation is visualised. It is evident that there is greater variance in the MCMC samples for Shrimp when compared to the Ladybird platform, particularly for the translation parameters, as corroborated by the numerical results in Table 3.

#### 4.4. Effect of Angular Diversity

To investigate how angular diversity of body poses in a dataset affects the certainty of the result, two experimental subsets were compiled from the outlier-rejected dataset with 16 observations for Ladybird. The first includes only ten observations with small roll angles  $\phi_{b,x}^w < 1.9^\circ$ . The second includes five observations with small roll  $\phi_{b,x}^w$  and five with roll angles  $3.3^\circ < \phi_{b,x}^w < 5.8^\circ$ . Both datasets therefore contain ten total observations, representing low and high angular diversity. Optimisation results for these subsets are shown in Table 4. Despite containing the same number of observations,

the dataset with high angular diversity results in significantly lower uncertainty for the optimal camera pose, compared to the low diversity set.



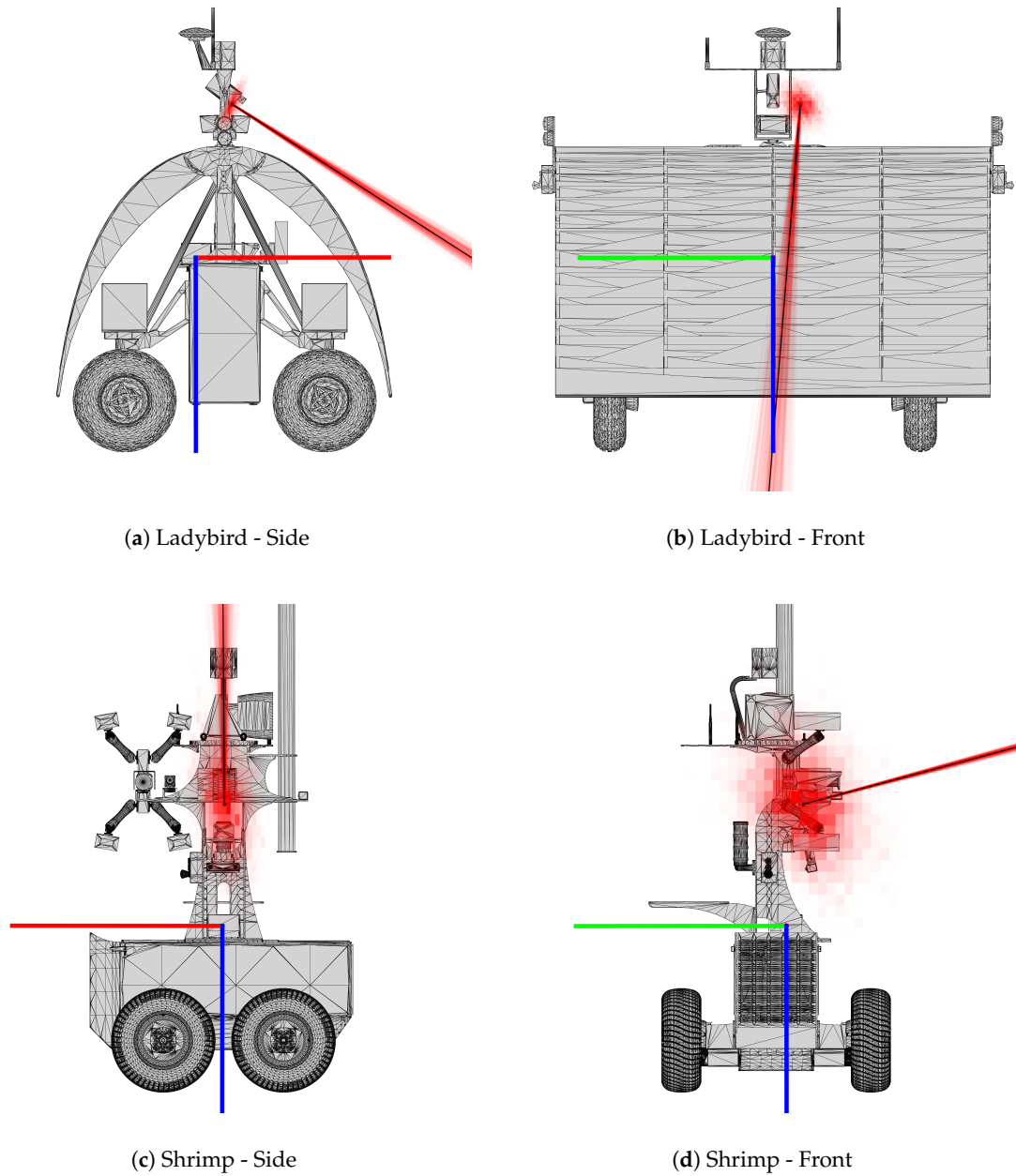
**Figure 11.** Ladybird and Shrimp platform MCMC axis-angle unit vector ( $\mathbf{e}_s^b$ ) samples plotted on a sphere (a,d) and close-up, centred to a pole (b,e) (using 16 and 14 observations for Ladybird and Shrimp respectively). Each point on the sphere is coloured according to a KDE to give an indication of the axis-angle vector marginal likelihood for visualisation purposes only. Points were randomly thinned by ten to facilitate plotting. A histogram of the magnitudes of the axis-angle rotation,  $\theta_s^b$ , is shown in (c,f). The hand measured pose orientation is shown with red crosses in (a,b,d,e), and a vertical red line in (c,f).

**Table 4.** Angular diversity comparison.

	$r_{c,x}^b$ (m)	$r_{c,y}^b$ (m)	$r_{c,z}^b$ (m)	$\theta_c^b e_{c,x}^b$ (rad)	$\theta_c^b e_{c,y}^b$ (rad)	$\theta_c^b e_{c,z}^b$ (rad)
High ang. diversity	$0.166 \pm 0.041$	$-0.155 \pm 0.083$	$-0.717 \pm 0.074$	$-0.831 \pm 0.023$	$0.744 \pm 0.023$	$-1.420 \pm 0.015$
Low ang. diversity	$0.147 \pm 0.118$	$0.058 \pm 0.107$	$-0.269 \pm 0.281$	$-0.797 \pm 0.051$	$0.796 \pm 0.059$	$-1.452 \pm 0.026$

#### 4.5. Repeatability

To assess the repeatability of the approach, another two experimental subsets were compiled from the outlier-rejected dataset with 16 observations for Ladybird. They each contained five observations with small roll  $\phi_{b,x}^w$  and three with roll angles  $3.3^\circ < \phi_{b,x}^w < 5.8^\circ$ . The two subsets do not share any observations, making them independent. Optimisation results for these subsets are shown in Table 5. The two results are consistent with each other, as the two distributions overlap to a significant degree. Due to the small dataset size, uncertainty values are higher than the result in Table 3 with all 16 observations, which is expected as demonstrated in Section 4.3.



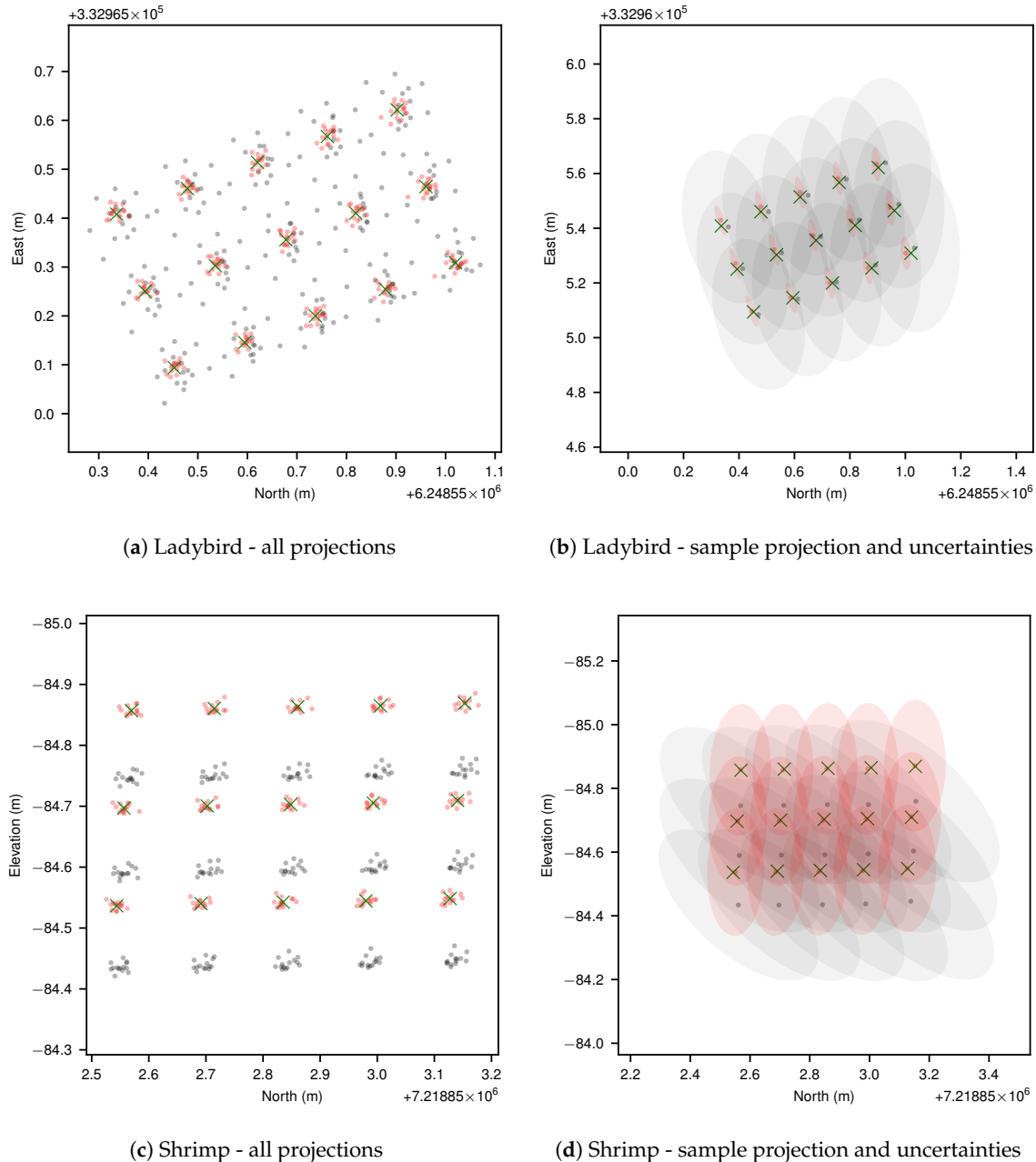
**Figure 12.** Optimised camera pose and MCMC sample poses for Ladybird (a,b) and Shrimp (c,d). The optimised pose is shown as an  $xyz$  axis with black lines. The MCMC samples are shown as red, green and blue  $xyz$  axes, where greater colour intensity corresponds to greater sample density, approximating the marginal likelihood of the poses in the viewing direction. The body (i.e., navigation system) pose is also shown as a red, green and blue  $xyz$  axis.

**Table 5.** Repeatability.

	$r_{c,x}^b$ (m)	$r_{c,y}^b$ (m)	$r_{c,z}^b$ (m)	$\theta_c^b e_{c,x}^b$ (rad)	$\theta_c^b e_{c,y}^b$ (rad)	$\theta_c^b e_{c,z}^b$ (rad)
Subset 1	$0.192 \pm 0.046$	$-0.170 \pm 0.095$	$-0.702 \pm 0.093$	$-0.828 \pm 0.024$	$0.735 \pm 0.028$	$-1.420 \pm 0.015$
Subset 2	$0.206 \pm 0.071$	$-0.306 \pm 0.123$	$-0.819 \pm 0.112$	$-0.839 \pm 0.034$	$0.693 \pm 0.040$	$-1.408 \pm 0.022$

#### 4.6. Mapping Accuracy

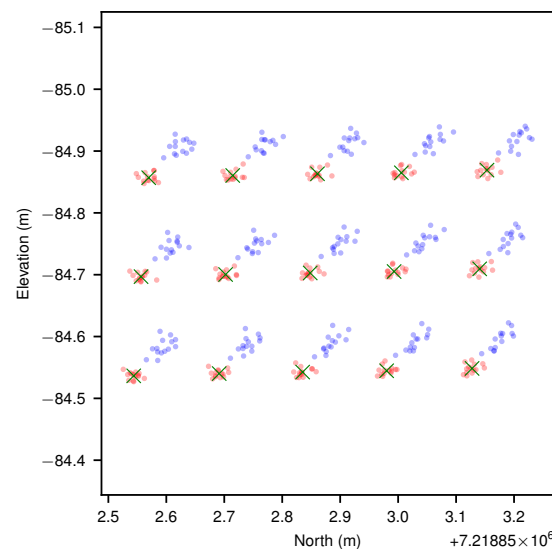
Figure 13 shows Ladybird's and Shrimp's observations, after outlier removal, projected to the best fit plane of predicted point locations before and after calibration (see Section 3.6). Figure 13a,c give projections of all observations, while Figure 13b,d only show the projections for one observation, but include uncertainty ellipses. Post calibration average pattern point estimates are also shown as green crosses.



**Figure 13.** Point to plane projection comparisons for both platforms. In (a,c) all scans of all pattern points are plotted to best fit planes before (grey) and after (red) calibration, while in (b,d) a single projected observation of each pattern point is shown with  $1\text{-}\sigma$  uncertainty ellipses. In all figures, post-calibration pattern point estimates are marked with green crosses for reference. Because the calibration pattern was positioned flat on the ground for Ladybird, a top down (North-East) view was selected for (a,b). Conversely, for Shrimp the calibration pattern was positioned almost vertically on a ladder, facing west, and therefore a side view (North-Elevation) was selected for (c,d).

For both platforms, the calibrated camera pose exhibited more densely bound projected points. The change in spread is less pronounced for the Shrimp platform, because as mentioned previously the manual measurements were by chance much closer to the optimum values, though a significant offset can be observed between manual and calibrated results. The plots also demonstrate the effect of relative camera pose uncertainty on mapping uncertainty, which were significant for both Ladybird and Shrimp.

Given the hand measurement for shrimp happened to be close to the optimal, the effect of adding just one degree of error to the measured axis-angle vector is shown in Figure 14, which compares the mapped points from the optimised and erroneous camera pose. The optimisation improves the cluster significantly compared to a hand measured pose with one degree error.

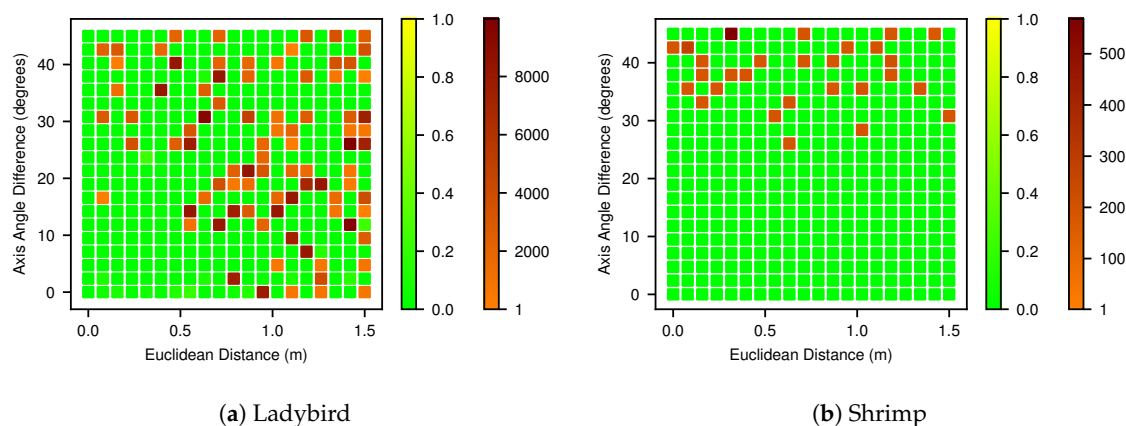


**Figure 14.** All pattern points plotted to best fit planes for Shrimp, red for post calibration (same as in Figure 13c) and blue for projected points resulting from a relative camera pose where the optimised orientation was altered by  $1^\circ$ . This demonstrates the significant effect small changes in the camera pose can have on mapping performance.

#### 4.7. Basin of Attraction

Basin of attraction plots, which were generated as described in Section 3.8 using the Powell optimiser, are shown in Figure 15 for both platforms. The Mahalanobis distances were generally either close to zero or very large, so they are colour coded into two tiers, below and above 1.0, to improve readability. The basin for the Ladybird platform (Figure 15a) shows success can be expected in the triangular region with less than  $20^\circ$  and 0.5 m of hand measurement error. The basin for Shrimp (Figure 15b) shows a greater immunity to translation errors and successful results can be expected with initialisation errors less than  $20^\circ$  and as high as 1.5 m. Both figures indicate that when the initialisation error is higher, there is still a high chance (approx. 60%) of a solution within a Mahalanobis distance of 0.1, but it cannot be relied upon, and deteriorates as the distance from optimum increases.





**Figure 15.** Basins of attractions for the (a) Ladybird and (b) Shrimp platforms. The x and y axes of the plots are the Euclidean distances and axis-angle rotational differences between the initial values and the optimal reference solution respectively. Each grid cell is colour coded into two tiers based on the Mahalanobis distance of the result to the known optimum: below 0.1 and above, which was chosen as a suitable threshold for optimisation success.

## 5. Discussion

The results show that the proposed method was able to reliably estimate the relative line scan camera pose on a mobile ground vehicle, resulting in a reduction in mapping error, as long as the calibration data includes sufficient viewpoint variability. An uncertainty of 0.06 m/0.018 rad (1.05°) for Ladybird, and 0.18 m/0.042 rad (2.39°) for Shrimp was achieved. This result is dependent on the certainty of input parameters, which include pixel observations, navigation system solutions, and camera intrinsics. For example, confidence in the calibrated pose parameters for the Ladybird platform was significantly greater than for the Shrimp platform, due to Ladybird's higher grade IMU, which allowed the navigation system to provide more certain solutions.

The results show that it is necessary to examine reprojection errors and remove outliers, as is common with many camera calibration approaches. Outliers statistically fall outside the assumptions encoded in their respective error models, and so the mean of the final camera pose distribution is pulled in the wrong direction. A number of observations for both platforms exhibited high reprojection errors relative to other observations (>approx. 8 pixels). These errors could be caused by manual labelling inaccuracies (e.g., due to limited resolution), navigation system solution errors that incorrectly fall outside the reported navigation uncertainty, or a combination thereof. Removing outliers had significant effects on the results, evident particularly in the correction of some parameters such as the z-offset  $t_{c,z}^b$  for both platforms (see Table 3). The results shown in Figure 8 also support the iterative removal of outliers at each stage. For instance, reprojection errors of Ladybird observation 20 improve as other observations are removed, while Shrimp observation 4 degrades, and was eventually removed. Conversely, as shown in Figure 9 a larger number of observations, outliers or not, allows for greater certainty in the final result. Thus, there are two competing factors when performing outlier rejection. A sufficiently large number of observations is required to maintain an acceptable level of certainty, yet removing outliers is important to minimise reprojection errors. It is, therefore, desirable to obtain a sufficient number of observations to allow for subsequent outlier removal. The additional computational time required when increasing the number of observations is also a consideration. The ending condition threshold should be chosen such that significant outliers are removed, but a sufficient number of observations remain. In this paper a value of 5 pixels was empirically determined as an appropriately balanced threshold given the data.

The main product of the MCMC uncertainty analysis is a covariance matrix (Equation (23)), which can be used to estimate mapping accuracy (e.g., Figure 13a,c). However, covariance matrices represent

uncertainty in a compressed format, given the assumption that the likelihood function is normally distributed. The corner plots (Figure 10) provide a direct view of the MCMC sampling result, which qualitatively confirm that for both vehicles the normality assumption is justified: specifically that the distributions behave linearly within the sampled region, and the 1D histograms are qualitatively Gaussian in shape.

In this paper we propose methods of visualising sensor pose distributions in a human interpretable way, as depicted in Figures 11 and 12. The sphere plots and associated axis-angle magnitude histograms in Figure 11 present the same underlying data in the MCMC sample plots (Figure 10), but focus on human interpretability of the orientation parameters. The sphere provides a relatable reference that demonstrates how closely clustered the pose orientation is. Similarly, the visualisation in Figure 12 allows for human interpretation of the resulting camera pose and uncertainty (likelihood distribution) with respect to the platform models. These figures particularly highlight the greater uncertainty in translation parameters compared to orientation. They also confirm that the solutions are qualitatively “sensible” with respect to the physical platform models.

The primary objective of optimising the camera pose is to reduce mapping errors. This was demonstrated in the results by the tighter clustering of mapped calibration target points that was achieved post-calibration. The improvement was particularly noticeable for the Ladybird platform, and to a lesser extent for the Shrimp platform. By chance, the manually measured camera pose on Shrimp was much closer to the optimal result than it was for Ladybird, and so the mapping improvement for Shrimp was less pronounced. The camera location on the Shrimp platform was easier to access, due to the lower height and smaller footprint, compared to the Ladybird platform, which likely explains the better manual estimate. Nevertheless, such accurate manual measurements can typically not be guaranteed, and Figure 14 reveals the sensitivity of the map to small errors in camera pose, highlighting the need for calibration.

The results reinforce the importance of acquiring a calibration dataset that exhibits a wide variety of platform poses with respect to the calibration pattern. This was tested by optimising both with and without large body roll ( $\phi_{b,x}^w$ ) observations. Removing high  $\phi_{b,x}^w$  observations had a considerable effect, as shown in Table 4, where uncertainty approximately doubled and even tripled for some parameters.

The proposed method is able to deal with a wide range of initial hand measured values when paired with the Powell optimisation algorithm. In Figure 15 we propose an intuitive approach to visualising the basins of attraction, by reducing the 6 DOF initial parameter space to form a 2D pose-distance space, comprising Euclidean and axis-angle distance. The plots demonstrate that initial estimates that deviate up to 0.5 m or 20° are likely to result in a successful optimisation. Additionally, with even larger deviations there is still a better than even (60%) chance of success. However, this is highly dependent on the geometry of the sensor and platform, the acquired data and the chosen optimisation algorithm. In our case, the solution for the Shrimp platform was surprisingly robust to initial translation errors (up to 1.5 m), while some failures can be seen at over 0.5 m for the Ladybird platform. This may be the result of the greater platform roll and pitch angles (up to 17°) achieved with Shrimp during data acquisition. In addition, different optimisers will have varying abilities to deal with local minima and “flat” regions in the 6 DOF parameter space. Nevertheless, measurement error tolerances of  $\pm 0.5$  m and  $\pm 20^\circ$  should be practically achievable for most applications.

As shown in Table 5, the camera pose estimates obtained from two independent datasets for the same platform were consistent with each other. The results therefore verify that the proposed approach is repeatable by demonstrating that different sets of data from the same platform yield consistent results.

An important advantage of the proposed method is that exact dimensions of the calibration pattern do not need to be known. As such, a planarity assumption or assumptions about the distances or geometry between points are not required. This simplifies the method in the field because it is not affected by printing errors or damage/warping of the pattern which affects the relative geometry of the points. Furthermore, a calibration pattern could be manually produced in the field if necessary.

One important condition, however, is that individual points are uniquely distinguishable in the line scan image data.

## 6. Conclusions

This paper demonstrated a novel method for estimating a rigidly mounted line scanning camera's fixed 6 DOF pose relative to a mobile platform with a navigation system. The method is appropriate for ground or very low altitude applications, where the scene is relatively near the platform, as it does not require GCPs and uses a compact calibration pattern, the dimensions of which do not need to be known. Furthermore, it does not require data from auxiliary sensors such as full frame cameras. The approach involves imaging a calibration pattern with distinctly identifiable points from various platform poses, and using the navigation system and image data to triangulate their positions in the world frame. Reprojecting the points to the camera yields reprojection errors, which are used as a basis for outlier rejection, and then to calculate the likelihood given a candidate camera pose. By minimising the negative log likelihood, the optimal relative camera pose can be obtained. Given the likelihood function, an MCMC algorithm is able to estimate the certainty of the camera pose. The results demonstrate the effectiveness of the approach using two different mobile platforms with differing mounting configurations. The method was shown to be robust to relatively inaccurate initial hand measurements (within 0.5 m and 20°). Additionally, a number visualisations have been proposed to aid in human interpretation of the results. Future work will attempt to precisely specify platform pose requirements prior to data collection, automate and improve the pattern point extraction process, and explore the application of a robust optimisation routine or loss function to simplify the outlier rejection process.

**Acknowledgments:** This work is supported by the Australian Centre for Field Robotics (ACFR) at The University of Sydney and by funding from the Australian Government Department of Agriculture and Water Resources as part of its Rural R&D for profit programme. For more information about robots and systems for agriculture at the ACFR, please visit <http://sydney.edu.au/acfr/agriculture>.

**Author Contributions:** Both A.W. and J.U. gathered the relevant data in the field. A.W. conceived and implemented the method, while J.U. supervised the work and provided significant conceptual input. The paper was written by A.W. and reviewed by J.U.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Yang, C.; Everitt, J.H.; Bradford, J.M.; Murden, D. Airborne hyperspectral imagery and yield monitor data for mapping cotton yield variability. *Precis. Agric.* **2004**, *5*, 445–461.
2. Bethel, J.; Lee, C.; Landgrebe, D.A. Geometric registration and classification of hyperspectral airborne pushbroom data. *Int. Arch. Photogramm. Remote Sens.* **2000**, *33*, 183–190.
3. Lim, M.S.; Lim, J. Visual measurement of pile movements for the foundation work using a high-speed line-scan camera. *Pattern Recognit.* **2008**, *41*, 2025–2033.
4. Dale, L.M.; Thewis, A.; Boudry, C.; Rotar, I.; Dardenne, P.; Baeten, V.; Pierna, J.A.F. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Appl. Spectrosc. Rev.* **2013**, *48*, 142–159.
5. Pfaff, F.; Maier, G.; Aristov, M.; Noack, B.; Gruna, R.; Hanebeck, U.D.; Längle, T.; Beyerer, J.; Pieper, C.; Kruggel-Emden, H.; et al. Real-time motion prediction using the chromatic offset of line scan cameras. *at-Automatisierungstechnik* **2017**, *65*, 369–380.
6. Li, D.; Wen, G.; Qiu, S. Cross-ratio—Based line scan camera calibration using a planar pattern. *Opt. Eng.* **2016**, *55*, 14104.
7. Ramirez-Paredes, J.P.; Lary, D.J.; Gans, N.R. Low-altitude terrestrial spectroscopy from a pushbroom sensor. *J. Field Robot.* **2015**, *33*, 837–852.

8. Deery, D.; Jimenez-Berni, J.; Jones, H.; Sirault, X.; Furbank, R. Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* **2014**, *4*, 349–379.
9. Wendel, A.; Underwood, J. Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. In Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 16–21 May 2016; pp. 5128–5135.
10. Trierscheid, M.; Pellenz, J.; Paulus, D.; Balthasar, D. Hyperspectral imaging for victim detection with rescue robots. In Proceedings of the IEEE International Workshop on Safety, Security and Rescue Robotics (SSRR), Sendai, Japan, 21–24 October 2008; pp. 7–12.
11. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334.
12. Mostafa, M.M.R. Bore-sight calibration of integrated inertial/camera systems. In Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation (KIS 2001), Banff, AB, Canada, 5–8 June 2001; pp. 5–8.
13. Lobo, J.; Dias, J. Relative pose calibration between visual and inertial sensors. *Int. J. Robot. Res.* **2007**, *26*, 561–575.
14. Hol, J.D.; Schön, T.B.; Gustafsson, F. Modeling and calibration of inertial and vision sensors. *Int. J. Robot. Res.* **2010**, *29*, 231–244.
15. Dráréni, J.; Roy, S.; Sturm, P. Plane-based calibration for linear cameras. *Int. J. Comput. Vis.* **2011**, *91*, 146–156.
16. Hui, B.; Wen, G.; Zhao, Z.; Li, D. Line-scan camera calibration in close-range photogrammetry. *Opt. Eng.* **2012**, *51*, 53601–53602.
17. Li, D.; Wen, G.; Hui, B.W.; Qiu, S.; Wang, W. Cross-ratio invariant based line scan camera geometric calibration with static linear data. *Opt. Lasers Eng.* **2014**, *62*, 119–125.
18. Luna, C.; Mazo, M.; Lázaro, J.L.; Vázquez, J.F. Calibration of line-scan cameras. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 2185–2190.
19. Yao, M.; Zhao, Z.; Xu, B. Geometric calibration of line-scan camera using a planar pattern. *J. Electron. Imaging* **2014**, *23*, 13028.
20. Sun, B.; Zhu, J.; Yang, L.; Yang, S.; Niu, Z. Calibration of line-scan cameras for precision measurement. *Appl. Opt.* **2016**, *55*, 6836–6843.
21. Ma, W.; Bioucas-Dias, J.; Chan, T.; Gillis, N.; Gader, P.; Plaza, A.; Ambikapathi, A.; Chi, C. A signal processing perspective on hyperspectral unmixing: Insights from remote sensing. *IEEE Signal Process. Mag.* **2014**, *31*, 67–81.
22. Lenz, A.; Schilling, H.; Perpeet, D.; Wuttke, S.; Gross, W.; Middelmann, W. Automatic in-flight bore-sight calibration considering topography for hyperspectral pushbroom sensors. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2981–2984.
23. Perry, J.; Ahmed, M.; Abd-Elrahman, A.; Bowman, S.; Kaddoura, Y.; Watts, A. Precision directly georeferenced unmanned aerial remote sensing system: Performance evaluation. In Proceedings of the Institute of Navigation National Technical Meeting, San Diego, CA, USA, 28–30 January 2008; pp. 680–688.
24. Muller, R.; Lehner, M.; Muller, R.; Reinartz, P.; Schroeder, M.; Vollmer, B. A program for direct georeferencing of airborne and spaceborne line scanner images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2002**, *34*, 148–153.
25. Abd-Elrahman, A.; Sassi, N.; Wilkinson, B.; Dewitt, B. Georeferencing of mobile ground-based hyperspectral digital single-lens reflex imagery. *J. Appl. Remote Sens.* **2016**, *10*, 14002.
26. Wohlfeil, J.; Bucher, T. A modular, interactive software-concept for radiometric and geometric correction of airborne and spaceborne linescanner images. In Proceedings of the Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IX (SPIE), Berlin, Germany, 31 August–3 September 2009; Volume 7478.
27. Yeh, C.K.; Tsai, V.J.D. Self-calibrated direct geo-referencing of airborne pushbroom hyperspectral images. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Vancouver, BC, Canada, 24–29 July 2011; pp. 2881–2883.
28. Wohlfeil, J. Optical Orientation Determination for Airborne and Spaceborne Line Cameras. Ph.D. Thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Berlin, Germany, 2012.

29. Barbieux, K.; Constantin, D.; Merminod, B. Correction of airborne pushbroom images orientation using bundle adjustment of frame images. In Proceedings of the The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, Prague, Czech Republic, 12–19 July 2016; 2016 XXIII ISPRS Congress; Volume XLI-B3, pp. 813–818.
30. Habib, A.; Xiong, W.; He, F.; Yang, H.L.; Crawford, M. Improving orthorectification of UAV-based push-broom scanner imagery using derived orthophotos from frame cameras. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 262–276.
31. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: Cambridge, UK, 2003.
32. Berner, P. Technical Concepts: Orientation, Rotation, Velocity and Acceleration, and the SRM. Available online: <http://www.sedris.org/wg8home/Documents/WG80485.pdf> (accessed on 21 March 2017).
33. Underwood, J. *Reliable and Safe Autonomy for Ground Vehicles in Unstructured Environments*; University of Sydney: Sydney, Australia, 2009.
34. Huynh, D.Q. Metrics for 3D rotations: Comparison and analysis. *J. Math. Imaging Vis.* **2009**, *35*, 155–164.
35. Gellert, W.; Gottwald, S.; Hellwich, M.; Kästner, H.; Küstner, H. *VNR Concise Encyclopedia of Mathematics*, 2nd ed.; Van Nostrand Reinhold: New York, NY, USA, 1989.
36. Underwood, J.P.; Hill, A.; Peynot, T.; Scheduling, S.J. Error modeling and calibration of exteroceptive sensors for accurate mapping applications. *J. Field Robot.* **2010**, *27*, 2–20.
37. James, F. *Statistical Methods in Experimental Physics: 2nd Edition*; World Scientific: Singapore, 2006.
38. Foreman-Mackey, D.; Hogg, D.W.; Lang, D.; Goodman, J. Emcee: The MCMC Hammer. *Publ. Astron. Soc. Pac.* **2013**, *125*, 306–312.
39. MacKay, D.J. *Information Theory, Inference, and Learning Algorithms*; Cambridge University Press: Cambridge, UK, 2008.
40. Underwood, J.; Wendel, A.; Schofield, B.; McMurray, L.; Kimber, R. Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. *J. Field Robot.* **2017**, *34*, 1061–1083.
41. Underwood, J.P.; Calleja, M.; Taylor, Z.; Hung, C.; Nieto, J.; Fitch, R.; Sukkarieh, S. Real-time target detection and steerable spray for vegetable crops. In Proceedings of the International Conference on Robotics and Automation: Robotics in Agriculture Workshop, Seattle, WA, USA, 26–30 May 2015.
42. Wendel, A.; Underwood, J. Illumination compensation in ground based hyperspectral imaging. *ISPRS J. Photogramm. Remote Sens.* **2017**, *129*, 162–178.
43. Stein, M.; Bargoti, S.; Underwood, J. Image based mango fruit detection, localisation and yield estimation using multiple view geometry. *Sensors* **2016**, *16*, 1915.
44. Bargoti, S.; Underwood, J.P. Image segmentation for fruit detection and yield estimation in apple orchards. *J. Field Robot.* **2017**, doi:10.1002/rob.21699.
45. Underwood, J.P.; Hung, C.; Whelan, B.; Sukkarieh, S. Mapping almond orchard canopy volume, flowers, fruit and yield using LiDAR and vision sensors. *Comput. Electron. Agric.* **2016**, *130*, 83–96.
46. Powell, M.J.D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* **1964**, *7*, 155–162.
47. Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open source scientific tools for Python.
48. Foreman-Mackey, D. corner.py: Scatterplot matrices in Python. *J. Open Source Softw.* **2016**, *24*, doi:10.21105/joss.00024.

