2008 IEEE International Conference on
Robotics and Automation
Pasadena, CA, USA, May 19-23, 2008

Removing Scale Biases and Ambiguity from 6DoF Monocular SLAM Using Inertial

Todd Lupton and Salah Sukkarieh

Abstract— **This paper identifies various scale factor biases commonly introduced into monocular SLAM implementations as a result of the true scale factor of the map not being observable. A way to make the scale factor observable and remove any scale biases via the use of an inertial measurement unit (IMU) is presented and implemented. Results show that with an IMU the true scale of the map becomes observable over time and the use of a square root information filter allows the effect of initial scale biases to be removed completely from the solution resulting in an unbiased solution no matter what the initial scale assumptions are.**

I. INTRODUCTION

One problem with using a single camera for SLAM is that due to its projective nature the scale of the environment in which it operates is not observable. There have been many successful implementations of monocular SLAM such as [1] where a set of initial base landmarks at a known location are used to make the map scale observable and constrain the estimate of the camera location and orientation until new landmarks can be initialized, and [2] where initially know landmarks are not required as an undelayed landmark initialization technique is used. Despite the success of these implementations problems with scale observability and consistency still remain.

In [2] even though no initial landmark locations are known and no other sensors are used to provide map scale information a full map and camera trajectory is recovered which is counter intuitive as the scale should not be observable. The scale information for this implementation comes from two false sources. These will be referred to as sources of scale bias as even when true scale information is available they bias the scale of the map.

These two sources of scale bias are:

1) The assumption of the variance of the camera's acceleration.
2) The initial range estimate to landmarks.

The assumption of the camera's acceleration variance is a byproduct of using a constant velocity model with no observations of the platform's true acceleration. The initial range estimate to landmarks is part of the undelayed inverse depth landmark initialization technique presented in [3] and used in [2]. An initial range guess is assigned to landmarks to use for linearization purposes and to make the solution full rank even though this information is not available at the time of initialization, or ever in [2]. The assumption of the camera's acceleration variance is also present in the implementation given in [1].

Problems are also caused when two different sources of scale bias are used and provide conflicting information about the scale of the map or if a single source of scale bias provides conflicting information about the scale in different parts of the map. An example of this is where an initial range estimate is used and in one region of the map the majority of the features are very close to the camera resulting in a map scale that is larger than the true scale and then the camera moves to another area where the majority of the features are far away from the camera resulting in a map scale that is smaller than the true scale. This situation will produce a map that has different scales in different regions. Also if the constant velocity model acceleration variance assumption suggests a different scale to the initial landmark range assumption, larger than expected observation innovations will result as well as suboptimal or overconfident updates.

In [4] the authors used a low cost inertial measurement unit (IMU) with monocular SLAM to remove the acceleration variance assumption and attempt to make the true map scale observable. However the initial range estimate bias remained along with the linerizations made based on those initial estimates. In most cases the initial range bias estimate will be the dominant source of scale bias especially when low cost IMUs are used.

This paper presents a technique that allows monocular SLAM with a single camera and a low cost IMU to be performed where true map scale becomes observable and any scale biases are removed. A square root implementation of a delayed state information filter [5] that has been adapted for use with a 6 degree of freedom (6DoF) inertial process model is used to allow initial range estimates to be removed and observations relinearized once sufficient information has been obtained about a landmark's true range.

II. METHODOLOGY

A. Inertial Measurement Units and Inertial Process Models

An IMU is a useful sensor for navigation and SLAM as it provides information about the motion of the platform it is attached to that is independent of the characteristics of the platform and does not require any external infrastructure or information. The IMU measures the linear acceleration and rotation rates of the platform which can be integrated to provide estimates of orientation, velocity and position with respect to an inertial frame

Todd Lupton and Salah Sukkarieh are with the ARC Centre for Excellence in Autonomous Systems, Australian Centre for Field Robotics, University of Sydney, Australia {t.lupton, salah}@cas.edu.au

The disadvantage of IMUs, especially low cost units are that the measurements obtained contain biases that must be estimated and removed before integration. The presence of noise in the measurements coupled with the integration to provide orientation, velocity and position estimates also results in the estimates drifting over time and external observations must be used to constrain this drift. Initial orientation, position and velocity estimates are also required as a starting point for integration.

B. Square Root Information Smoothing

The representation of the SLAM problem in information space rather than the more common state space has received some interest recently [5][6][7][8]. The information form of a Gaussian distribution is often called the canonical form and can be represented by an information matrix, $\mathbf{Y}$, and information vector, $\mathbf{y}$, that can be derived from the mean, $\mathbf{x}$, and covariance, $\mathbf{P}$ of the state space distribution by:

$$\mathbf{Y} = \mathbf{P}^{-1} \qquad \mathbf{y} = \mathbf{Y}\mathbf{x} \tag{1}$$

Of particular interest to this application is the delayed state (smoothing) square root information form presented in [5] as it allows observations to be stored separately in the square root information matrix, $\mathbf{A}$, where:

$$\mathbf{Y} = \mathbf{A}'\mathbf{A} \tag{2}$$

and process model predictions and landmark observations are simply included as extra independent rows of the square root information matrix.

C. State Parameterization

It is common in inertial SLAM to parameterize the vehicle state as a position, velocity and attitude ([9]) as these values are required to perform the integration of inertial observations. Since a delayed state filter is being used in this case all past positions are maintained in the filter which makes the velocity component of the vehicle state redundant as it can be recovered from two subsequent positions. Therefore the vehicle state has been reparameterized to contain only the position and attitude estimates of the IMU as can be seen in (3).

$$\mathbf{x}_t = \begin{bmatrix} pos_t \\ att_t \\ pos_{t-1} \\ att_{t-1} \\ pos_{t-2} \\ att_{t-2} \\ \vdots \\ accn_{bias} \\ gyro_{bias} \end{bmatrix} \tag{3}$$

This reparameterization is also required to keep the information matrix full rank and therefore allow an unconstrained solve to be used to recover an estimate of the state means. Reducing the number of elements in the information vector reduces the computational complexity as well.

D. Prediction and the Process Model

Normal information smoothing prediction is derived from the Kalman filter prediction equations and is performed as:

$$\begin{aligned} \mathbf{Y}_{t+1} = &\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_{v_t v_t} & \mathbf{Y}_{v_t m} \\ \mathbf{0} & \mathbf{Y}_{m v_t} & \mathbf{Y}_{mm} \end{bmatrix} \\ &+ \begin{bmatrix} \left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1} & -\left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1}\mathbf{F} & \mathbf{0} \\ -\mathbf{F}'\left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1} & \mathbf{F}'\left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1}\mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{aligned} \tag{4}$$

The first matrix in (4) contains the old information matrix. The $\mathbf{Y}_{v_t v_t}$ term relates to the the last two vehicle positions that are used in the prediction, and the $\mathbf{Y}_{mm}$ term relates to all other past vehicle positions and all the landmarks. The second matrix contains the terms relating to the new predicted position and the previous positions that this new position was calculated from. It can be seen that there are only off diagonal terms between the new current position and the last vehicle position and all the terms for the previous vehicle positions and landmark terms are left unchanged. This results in a constant time prediction in contrast to the cubic time required for prediction in the information filtering approach, due to the required marginalising out of old position states, and the linear time required for Kalman filter predictions, due to the required updating of the cross-covariance terms between the new vehicle position states and all the landmark states in the filter.

It can be seen from the above equation that $\left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1}$, which is the increase in vehicle position uncertainty due to the process model noise, is required to be invertible. Reparameterization of the vehicle state from standard models may be necessary to satisfy this requirement.

The prediction equations for the square root form of information smoothing can be easily derived from (4) as:

$$\mathbf{A}_{t+1} = \begin{bmatrix} \sqrt{\left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1}} & -\sqrt{\left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1}}\mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{v_t} & \mathbf{A}_m \end{bmatrix} \tag{5}$$

Equation (5) shows the advantage of this particular square root form in keeping predictions (and as will be seen later, observations as well) separate from each other which allows for easy relinearization at a later time. The very sparse nature of the information matrix also leads to computational advantages.

During a prediction, elements relating to the new vehicle position states are appended to the top of the information vector, $\mathbf{y}$, and the elements relating to the last two vehicle positions, which the new position has non-zero off diagonal terms with in the new information vector are updated. This process can be seen in (6) below:

$$\mathbf{y}_{t+1} = \begin{bmatrix} \mathbf{0} \\ \mathbf{y}_{vt} \\ \mathbf{y}_{mt} \end{bmatrix} + \begin{bmatrix} \mathbf{I}' \\ -\mathbf{F}' \\ \mathbf{0} \end{bmatrix} \left(\mathbf{G}\mathbf{Q}\mathbf{G}'\right)^{-1} \left(\mathbf{0} - \left(f\left(\mathbf{x}_t\right) - \mathbf{F}\mathbf{x}_t\right)\right) \tag{6}$$

The $\left(f(\mathbf{x}_t) - \mathbf{F}\mathbf{x}_t\right)$ term in the equation represents the linearization error for the process model. As the vehicle

state has been reparameterized to remove the vehicle velocity estimates the process model used has to be modified from the standard inertial process model ([9]) to the second order Markov model shown in (7).

$$f(\mathbf{x_t}) = \begin{bmatrix} pos_{t+1} \\ att_{t+1} \end{bmatrix} = \begin{bmatrix} pos_t + \left(\frac{pos_t + pos_{t-1}}{dt} + \left(\mathbf{C}_b^n \left(\mathbf{f}_b + acc_{bias}^n\right) + \mathbf{g}\right) dt\right) dt \\ att_t + \mathbf{E}_b^n \left(\omega_b + gyro_{bias}\right) dt \end{bmatrix} \quad (7)$$

Where dt is the time between predictions, $\mathbf{f}_b$ is the accelerometer observations, ω_b is the gyro observations, $\mathbf{C}_b^n$ is the rotation matrix from the IMU body frame to the navigation frame and $\mathbf{E}_b^n$ is the rotation rate transformation matrix between the body frame and the navigation frame.

This results in process model Jacobian matrices of

$$\mathbf{F} = \frac{f(\mathbf{x}_t)}{dx} = \begin{bmatrix} 2\mathbf{I} & \frac{d\mathbf{C}_b^n \mathbf{f}_b}{datt} dt^2 & -\mathbf{I} & \mathbf{0} & \dots & \mathbf{C}_b^n \mathbf{f}_b dt^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} + \frac{d\mathbf{E}_b^n \omega_b}{datt} dt & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{E}_b^n dt \end{bmatrix} \quad (8)$$

and

$$\mathbf{G} = \frac{f(\mathbf{x}_t)}{du} = \begin{bmatrix} \mathbf{C}_b^n dt^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_b^n dt \end{bmatrix} \quad (9)$$

E. Undelayed Feature Initalisation and the Observation Model

A simple pin-hole camera model is used after undistortion is performed on the images, this provides an azimuth and elevation bearing of observed landmarks which is used in the filter.

The inverse depth undelayed initialization technique presented in [3] is used as this application is in an unstructured environment where there are no initially know landmarks to start from and as the IMU is low cost, landmarks need to be used as soon as possible after first observation to constrain the drift of the IMU predictions. Therefore the landmarks are stored as azimuth, elevation and inverse range from the camera's position on first observation, the position states for the camera at this time do not need to be stored with the feature states as they are already present in the smoother, so only an index is stored externally to indicate which position the feature was first observed from.

Incorporating landmark observations into the square root information smoother is done in a similar way to performing a prediction. An observation update on a standard form information matrix in an information smoother, as presented in [6], is performed as:

$$\mathbf{Y}_{t|t} = \mathbf{Y}_{t|t-1} + \mathbf{H}'\mathbf{R}^{-1}\mathbf{H} \quad (10)$$

where $\mathbf{H}$ is the observation model Jacobian and $\mathbf{R}$ is the sensor observation noise covariance matrix. From this the square root information matrix observation update can be derived from (10) as:

$$\mathbf{A}_{t|t} = \begin{bmatrix} \mathbf{A}_{t|t-1} \\ \mathbf{v}^{-1}\mathbf{H} \end{bmatrix} \quad (11)$$

where:

$$\mathbf{v}'\mathbf{v} = \mathbf{R}$$

As can be seen from (11) the observation update of the square root information matrix only involves appending an additional row to the matrix which keeps this observation separate from all other operations allowing it to be easily modified at a later time if desired.

The information vector, $\mathbf{y}$, can be updated as in a normal information smoother as:

$$\mathbf{y}_{t|t} = \mathbf{y}_{t|t-1} + \mathbf{H}'\mathbf{R}^{-1}\left(\mathbf{z}_t - \left(h\left(\mathbf{x}_t\right) - \mathbf{H}\mathbf{x}_t\right)\right) \quad (12)$$

where $\mathbf{z}$ is the observation and $(h(\mathbf{x}_t) - \mathbf{H}\mathbf{x}_t)$ is the linearization error of the observation function.

The initial range estimate from the first observation of a landmark is simply treated as an observation of the range to the landmark and processed as any other observation.

F. State Recovery

An estimate of the state means is required as a linearization point for the observation and prediction functions and this is performed incrementally using preconditioned conjugate gradients. This technique is efficient as a fairly accurate estimate of all the state means is available at all times from the previous iteration of the smoother allowing for sufficient convergence in a small number of iterations.

G. Removing Initial Range Estimates and Relinearizeing

After sufficient accelerations have been observed so that the scale of the map becomes observable it is possible to remove some or all of the initial range estimates of landmarks to remove the biasing effect they have on the final solution. In the particular square root information form used this simply involves removing the row of the square root information matrix, $\mathbf{A}$, corresponding to this initial range estimate for a landmark, no other operations need to be performed.

After removing an initial range estimate or just as the filter runs and more information is obtained a more accurate estimate of landmark and past camera locations is obtained. This allows more accurate linearization points to be calculated and therefore it is beneficial to reperform the linearization of past observation and prediction functions at this time and resolve for state means if they are required.

III. EXPERIMENTAL SETUP

A. Hardware

The hardware used in this experiment is a Fire-i grayscale OEM firewire camera and a MicroStrain 3DM-GX1 inertial measurement unit. These sensors are mounted together on a fire fighting helmet so that they can be used to map the motion of the person wearing it.

Fig. 1. Sample images taken from the data set used to generate the presented results. Images were acquired at 7.5 fps and features were extracted using a Harris corner detector and matched using optical flow. The feature id numbers and feature motion from the previous image is shown.

B. Experiment

As a low cost IMU is being used and it is desirable for scale to become observable as soon as possible to test the algorithm, a trajectory that consists of repeated side to side motion of the camera over a range of about 1 meter while looking in the same direction was used. This trajectory allows regular repeated periods of linear acceleration, as well as establishing a base line for landmarks quickly so that both scale and range to landmarks is observable. Sample images from the dataset used can be seen in Fig. 1.

Feature extraction is performed by using a Harris corner detector and matching is done by optical flow as covariance information is not easily obtainable when information filters are used. This feature matching technique is also independent of the filter estimate so that a bad filter estimate will not result in continually incorrect data associations. While more complex feature extraction and matching algorithms could be beneficial to this implementation, such as correlation matching or SIFT features, this approach works sufficiently well on the datasets used and as feature matching is not the focus of this paper optical flow was deemed sufficient for our purposes.

Images were acquired at 7.5 frames per second as this was sufficient for this experiment and the feature matching method used. The IMU was run at approximately 152 samples per second.

This filter is implemented as a square root information filter as described in section II, this form was chosen to allow removal of initial range estimates and relineraization. A naive implementation is used and therefore real time performance is not achieved but the authors believe with the many optimization techniques available this will be achievable in the future.

IV. RESULTS

A. Initial Result

The information smoother was run as described over 64 images (approximately 8.5 seconds) with an initial estimated range to new landmarks of 1 meter. As in this dataset the camera was held 1 meter from the desk shown this initial range estimate is approximately correct for the features observed.

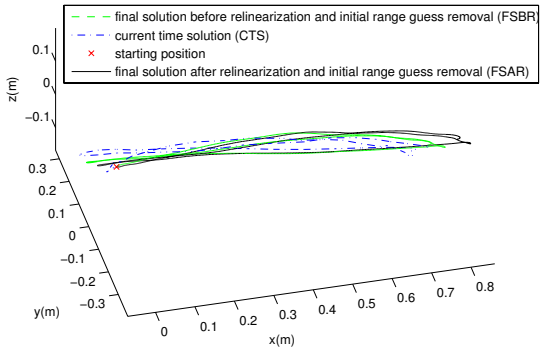

Fig. 2. Estimated camera path with a 1m initial range estimate before and after removal of initial range estimates and relinearization. The current time solution is the solution available when the camera is at that location, this is the same as the extended Kalman filter solution as only past information is used and no relinearization is performed.

The three different estimated camera trajectories produced by the filter on this dataset can be seen in Fig. 2. These different camera trajectory estimates are produced in different stages of the filters estimate and they are:

1) The current time solution (CTS).
2) The final solution before relinearization and initial range guess removal (FSBR).
3) The final solution after relinearization and initial range guess removal (FSAR).

The current time solution (CTS) of the camera position is the estimated camera position when the camera was at that location and therefore it only uses information gathered from observations before that point in time. As this estimate only uses past information and relinearization has not yet been performed it is the same as the estimate obtained from an extended Kalman filter (EKF). The final solution before relinearization and initial range guess removal (FSBR) is the estimate of the cameras path using all information gathered over the entire run but before removal of the initial range estimates and relinearization is performed. The final solution after relinearization and initial guess range removal (FSAR) is the estimate of the cameras path using all information gathered over the entire run after the initial range estimates have been removed and relinearisation of all observation and

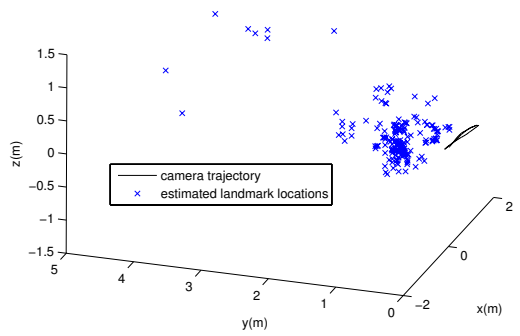

Fig. 3. Estimated landmark locations and camera path with a 1m initial range estimate after removal of initial range estimate and relinearization.

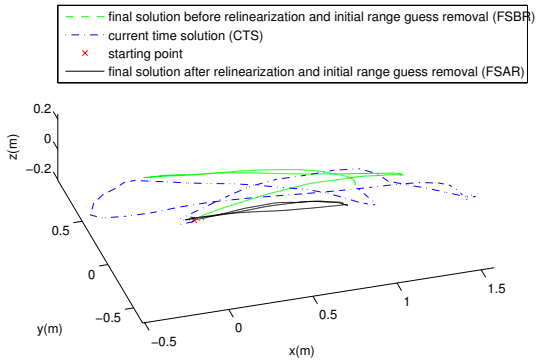

Fig. 4. Estimated camera path with a 10m initial range estimate before and after removal of initial range estimates and relinearization. The current time solution is the solution available when the camera is at that location, this is the same as the extended Kalman filter solution as only past information is used and no relinearization is performed.

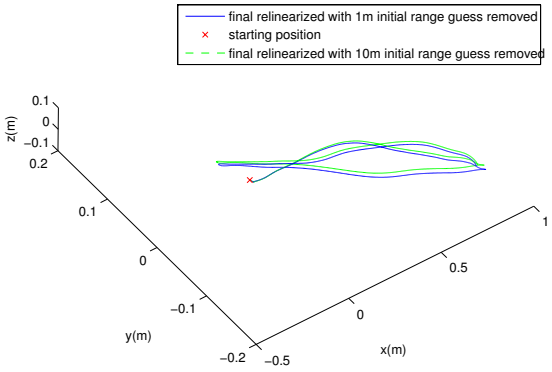

Fig. 5. Comparison of the final estimated trajectory after removal of initial range estimates and relinearization for initial range estimates of 1m and 10m. It can be seen from this figure that the solutions are almost identical and therefore independent of the initial range estimates used.

prediction functions has occurred so that no scale biasing information remains in the solution. This is the result of interest in this paper.

It can be seen from Fig. 2 that the FSAR is very similar to the FSBR. This is because the initial range estimates are fairly close to the true ranges to landmarks in this case but this will not be the case if the initial range estimates are far from the true ranges as will be seen in the next subsection.

The CTS of the camera trajectory differs much more from the final smoothed estimate especially in the beginning when the range to the landmarks as well as the IMU orientation and biases are not well known, this is where the bulk of the linearization errors will occur. It can be seen in the beginning of the run that the estimated direction of motion for the camera is actually in the wrong direction. This can have disastrous consequences for range initialization and feature matching if the filter estimates are used for assessing data association. The initial wrong direction for the estimated motion is a result of the IMU biases and a very coarse initial estimate of the IMU orientation.

A full map is produced with estimates of the locations of observed landmarks. The final map estimate along with the estimated camera trajectory after removal of initial range estimates and relinearization can be seen in Fig. 3. It can be seen from this graph that the majority of the landmarks have a final range estimate the is approximately 1 meter from the camera's trajectory which is expected for this dataset.

B. Removing Initial Range Estimates and Relinearizing

Fig. 4 shows an example of what happens if the initial range estimate to features is significantly different from the truth. This CTS result also shows what can be expected if two different sources of false scale bias give conflicting scale information, as in this case for the CTS estimate the accelerometers and the initial range guesses give different estimates of the map and trajectory scale.

It can be seen that the CTS estimate of the camera location is very unstable as a result of the IMU and feature observations giving conflicting information about the motion of the camera. Even the FSBR in this example shows a trajectory that is significantly different from the truth in both scale and relative motion. However, once the initial range estimates are removed and the whole trajectory is relinearized it can be seen that the FSAR solution converges to something resembling the true path of the camera in both scale and relative motion and this is directly comparable to the FSAR solution for initial range estimates of 1m as can be seen in Fig. 5. This is as expected as with the IMU and the removal of the initial range estimates and relinearizaton the solution no longer contains any heuristic scale biasing information and should converge to the same solution no matter what the initial range estimate was.

C. Scale Observability and Numerical Conditioning

As the camera is a projective sensor and cannot observe scale, the only source of scale information available to the smoother is from linear accelerations sensed by the IMU. As a low cost IMU was used for this experiment the scale may

only be weakly observable. The trajectory used was selected as it will give the best possible chance of the scale being observable.

If the absolute range to landmarks and therefore the map and trajectory scale was not observable, when the initial range estimates are removed the information matrix would become singular as it would not contain enough information to determine all the means in the state vector. We can use the condition number of the final information matrix as a measure of how close the matrix is to being singular which can indicate how well observed the scale of the map and trajectory is.

Fig. 6 shows a comparison of the condition numbers of the information matrices before and after removal of the initial range estimates for different lengths of time that the smoother was running. From this figure we can see how the condition number of the information matrices without the initial range estimates for only a few processed images is higher, indicating that they are close to being singular and therefore the scale of the map is only weakly observable. As the number of processed images increases the condition number of this information matrix decreases as the scale becomes more and more observable. The flattening out of the condition numbers at the end is considered to be a result of the fact that optical flow was used for data association in this example and if features were lost they are discarded and new ones obtained, if the same features were tracked throughout the entire run then it would be expected that this condition number would continue to decrease. A different feature matching algorithm, such as the ones discussed in section III, could be used to achieve this.

The condition number for the information matrices containing the initial range information can be seen to be significantly lower than the equivalent matrices without initial range estimates, which is as expected as they contain false information about the range to features making it appear that the map scale is observable even though it may not be. If a higher quality IMU was used the condition numbers for the case without the initial range estimates would be closer to the case with initial range estimates as the true scale would be more observable. In fact for the dataset used the average accelerations experienced are close to or below the noise floor of the IMU so it does not add much to the process model prediction accuracy. A constant velocity model may even make more accurate predictions in this case, however the IMU is still able to make the true map scale observable over a number of observations.

V. CONCLUSIONS

When performing SLAM with a single camera as the only sensor, implicit biasing of the scale of the map through assumptions of acceleration noise in the constant velocity model and/or explicit biasing from initial landmark range estimates can have a detrimental effect on the quality of the solution, especially when the scale assumptions are wrong.

Using an IMU in conjunction with the camera allows the true camera accelerations to be observed. Even if a low cost IMU is used this will avoid the biasing of the map scale from assuming an acceleration variance, even if its quality is too low to add to the accuracy of the process model predictions.

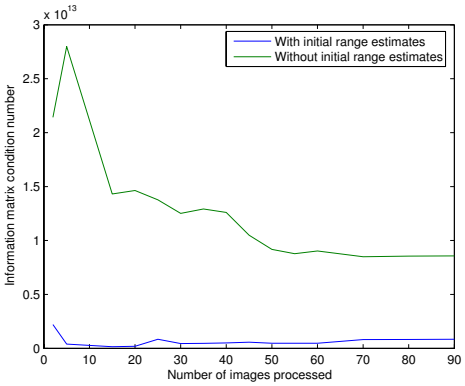

Fig. 6. Condition number of the information matrix with and without the initial range estimates for different numbers of images processed. This number can be used as a measure of the observability of the solution and it can be seen from the figure that the scale of the map without initial range estimates becomes more and more observable over time as more accelerations are observed.

Furthermore over time as linear accelerations are observed and combined with the visual feature observations the true scale of the map will become observable, resulting in a feature map and vehicle trajectory estimate that is free from heuristic assumptions about scale.

This paper has presented a method that allows a low cost IMU to be used in monocular SLAM where scale biases are removed from the solution and the true map scale becomes observable.

VI. ACKNOWLEDGMENTS

This work is supported in part by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC) and the New South Wales State Government and by ARC Discovery Grant DP0665439.

REFERENCES

[1] A. Davison, "Real-time simultaneous localisation and mapping with a single camera," in *International Conference on Computer Vision*, Nice, October 2003.

[2] J. M. M. Montiel, J. Civera, and A. Davison, "Unified inverse depth parametrization for monocular slam," in *RSS*, Philadelphia, Pennsylvania, 2006.

[3] J. Sola, A. Monin, M. Devy, and T. Lemair, "Undelayed initialization in bearing only slam," in *IRS*, Edmonton, Canada, 2005, pp. 2499–2504.

[4] P. Pinies, T. Lupton, S. Sukkarieh, and J. D. Tardos, "Inertial aiding of inverse depth slam using a monocular camera," in *ICRA*, 2007.

[5] F. Dellaert and M. Kaess, "Square root sam: Simultaneous localization and mapping via square root information smoothing," *IJRR*, vol. 25, no. 12, pp. 1181–1203, 2006.

[6] R. M. Eustice, H. Singh, and J. J. Leonard, "Exactly sparse delayed-state filters for view-based slam," *IEEE Transactions on Robotics*, vol. 22, 2006.

[7] J. Folkesson and H. I. Christensen, "Graphical slam for outdoor applications," *Journal of Field Robotics*, vol. 24, pp. 51–70, 2007.

[8] Y. Lui and S. Thrun, "Results for outdoor-slam using sparse extended information filters," in *ICRA*, 2003.

[9] M. Bryson and S. Sukkarieh, "Bearing-only slam for an airborne vehicle," in *ACRA*, Sydney, Dec 2005.