Outlier Rejection for Visual Odometry using Parity Space Methods

Arun Das* and Steven L. Waslander†
University of Waterloo, Waterloo, ON, Canada, N2L 3G1

***Abstract*— Visual odometry is a useful tool for many applications in mobile robotics, however camera egomotion estimation can become imprecise when outliers are present within the observed measurements. Typically, random sample consensus (RANSAC) approaches are used to perform outlier rejection, however the use of RANSAC can become computationally expensive for visual odometry, as the system state and feature reprojection error must be computed at every iteration. The parity space approach provides methodology to perform computationally efficient consistency checks for observations, without having to explicitly compute the system state. This work presents two outlier based rejection techniques, Group Parity Outlier Rejection and Parity Initialized RANSAC which use the parity space approach to perform rapid outlier rejection. The proposed methods are applied to the LIBVISO2 visual odometry algorithm and are tested using the KITTI vision dataset. Experiments demonstrate the proposed approaches are able to compute solutions with increased accuracy and improved run-time when compared to RANSAC.**

I. Introduction

In situations where observed data is fitted to a parameterized model, the presence of outliers in the measurements can corrupt the solution. In mobile robotics applications such as visual odometry, localization, and SLAM, the inclusion of outlier data points in the estimation reduces the accuracy of the computed vehicle state, or in a worst case, results in a diverging solution.

Many approaches exist which aim to attenuate the effect outliers have on the solution, or remove them from the measurement set all-together. The use of robust statistical methods, such as the L-estimator, M-estimator, R-estimator [1] and the least median squares [2], were proposed in order to reduce the affect of outliers on the final solution. Although the robust statistical approaches are well studied, they aim to only attenuate the contributions of outliers, and do not completely remove them from the problem, and thus would have poor performance where the measurements contain a high outlier ratio. The Hough transform generates a set of parametrized models and has each measurement vote for the models with which it is compatible [3]. Assuming that the spurious measurements do not consistently vote for any single model, the Hough transform provides adequate outlier rejection performance. However, the approach performs a discretization of the parameter space, and thus has poor computational performance when a large number of parameters are to be estimated. More recently, a batch heterogeneous outlier detection algorithm was suggested and used within a SLAM framework [4], however the approach is mainly limited to scenarios where the number of observed measurements is small.

In order to perform outlier rejection where many measurements are observed, the RANSAC algorithm [5] is typically used, and has been successful in many computer vision [6], visual odometry [7], [8] and SLAM [9], [10] applications. Many extensions of the RANSAC algorithm have been proposed, such as Maximum Liklihood Sample Consensus (MLESAC) [11], which assumes known probability distributions of the inlier and outlier to evaluate the sample hypothesis, and Local Optimization RANSAC (LO-RANSAC) [12], where the maximum inlier set is refined at each iteration using a local optimization technique.

The general issue with the RANSAC approach is having to perform a sufficiently large number of iterations in order to achieve a model with a certain confidence level. The number of required iterations can become quite large when the outlier ratio is high, which is undesirable in situations where computing the model parameters or the inlier set at each iteration is computationally expensive. The Random-RANSAC (R-RANSAC) approach [13], [14] was suggested to reduce the number of iterations required. In the R-RANSAC approach, a preliminary test is performed on the random sample using a small subset of the full measurement data, and only once this test is successful is the random sample evaluated against the full set of observations. Although effective, the R-RANSAC approach still requires the calculation of the model parameters at each iteration.

A promising approach for outlier detection is the parity space approach (PSA), which was first developed in order to perform fault detection and isolation for instrument clusters [15], but has also been applied to fault detection for nuclear power stations [16], and in flight avionics [17]. The PSA has also been applied in a visual SLAM formulation [18] where batches of image features tracked over a sliding window are tested for outliers and spurious features are rejected using a parity space consistency test. Using the measurement model of the system, the PSA projects the measurements into the parity space where the presence of outliers can be detected. Since the projection onto parity space is computed using only the measurement model, the state of the system is removed from the problem, meaning the parameters of the state do not need to be calculated in order to perform outlier detection.

Visual odometry (VO) estimates the egomotion of the camera though examination of the changes that motion induces on the camera images. A typical VO approach

* PhD Student, Mechanical and Mechatronics Engineering, University of Waterloo; adas@uwaterloo.ca
† Assistant Professor, Mechanical and Mechatronics Engineering, University of Waterloo; stevenw@uwaterloo.ca

Preprint submitted to 2014 IEEE International Conference on Robotics and Automation. Received September 15, 2013.

is to use corresponding image features between successive camera frames and estimate the incremental motion, which can be successfully performed using both monocular [19], [20] and stereo cameras [8], [6], [21]. In order to provide an accurate motion estimation, feature correspondences should not contain outliers, and typically a rejection scheme such as RANSAC is used.

In this work, we propose two methods which use parity space consistency testing to perform outlier rejection and demonstrate their effectiveness when used in a visual odometry application. The first method, Group Parity Outlier Rejection (GPOR), a parity space test is performed on subgroups of the measurement vector, and groups which fail the test are discarded. Since the test occurs completely within the parity space, the computation of model parameters is not required, thus the approach allows for rapid outlier removal with computational complexity that scales linearly with the number of measurements.

In the second method, Parity Initialized RANSAC (PI-RANSAC), improves the RANSAC algorithm by performing a parity space consistency test on the randomly selected sample set. The parity space test is computationally efficient and does not require determination of the model parameters, and thus provides a rapid consistency test for the randomly selected sample set. The PI-RANSAC method is shown to generate accurate parameter estimates with significantly fewer iterations when compared to RANSAC.

The GPOR and PI-RANSAC methods are validated using using the KITTI vision dataset [22]. The experiments demonstrate that the GPOR approach consistently outperforms RANSAC in terms of run-time while providing a comparable accuracy in the solution, and the PI-RANSAC method is able to provide solutions with an average 44.45% increase in accuracy and an average 68.95% improvement in run-time when compared to RANSAC.

II. Problem Formulation

A. Visual Odometry Method

The visual odometry algorithm used in this work was initially developed by [7] and [8]. An open source implementation is available at `http://www.cvlibs.net/software/libviso/`. A brief overview of the approach is presented here.

To determine the egomotion of the stereo camera between two successive frames, feature detection and matching is performed between the images of the stereo pair at the previous time-step to the images of the stereo-pair at the current time-step. Denote the i^{th} 3D feature point from the previous time step as $d_i = [d_i^x \; d_i^y \; d_i^z]^T \in \mathbb{R}^3$, where d_i^x, d_i^y and d_i^z denote the x, y and z components, respectively, of the point in the camera frame. The translation between the left camera frame of the previous time-step to the left camera frame in the current time-step is given as t_x, t_y and t_z, while the rotation between frames are r_x, r_y and r_z, for rotation about the x, y and z axis, respectively. The full state which defines the ego-motion of the camera is denoted as $x = [t_x \; t_y \; t_z \; r_x \; r_y \; r_z]^T$. Let $R_x \in \mathbb{SO}(3)$, $R_y \in \mathbb{SO}(3)$ and $R_z \in \mathbb{SO}(3)$ be rotation matrices constructed from the parameters r_x, r_y and r_z, respectively. Using a 3-2-1 Euler parametrization, a transformation function $T_x(d) : \mathbb{R}^3 \mapsto \mathbb{R}^3$ is given as

$$T_x(d) = (R_x R_y R_z)d + [t_x \; t_y \; t_z]^T$$

Given N_Y features from the previous time-step, denote the full set of 3D feature points as $Y = \{d_1, \ldots, d_{N_Y}\}$. Provided an estimate of the egomotion is available, each feature from the previous time-step can be transformed according to the egomotion estimate and re-projected back into the pixel co-ordinates of the stereo images for the current time-step. Denote the left camera reprojection mapping, $\Theta_x(d)^l : \mathbb{R}^3 \mapsto \mathbb{R}^2$ as

$$\Theta_x(d)^l = \mathcal{K}(T_x(d))$$

where $\mathcal{K}$ is the camera matrix. Similary, a re-projection mapping into the right camera image is given as

$$\Theta_x(d)^r = \mathcal{K}(T_x(d) - [\gamma \; 0 \; 0]^T)$$

where γ is the baseline between left and right camera. Using the image pair from the current time step, a set of measurements to feature pixel locations which correspond to the features from the previous image can be generated. Denote a measurement of the feature corresponding to a point, $d_i \in Y$, in the current left image as $z^l_{d_i} \in \mathbb{R}^2$ and similarly a measurement in the current right image as $z^r_{d_i} \in \mathbb{R}^2$. Using the measured features of the current frame and the re-projected feature point locations from the previous frame, a cost function, $\Lambda_Y(x) : \mathbb{R}^6 \mapsto \mathbb{R}$, which penalizes the reprojection error over the entire set of corresponding features from feature set Y can be defined as

$$\Lambda_Y(x) = \sum_{d_i \in Y} \|z^l_{d_i} - \Theta_x(d_i)^l\|^2 + \|z^r_{d_i} - \Theta_x(d_i)^r\|^2 \quad (1)$$

Finally, the egomotion of the camera is estimated by optimizing the cost given by Equation 1 over the features contained in Y, and the optimal transformation parameters, x^*, between the previous and current frame is

$$x^* = \underset{x \in \mathbb{R}^6}{\operatorname{argmin}} \, \Lambda(x)_Y$$

In order to apply parity space techniques to the proposed visual odometry method, the full measurement model must also be defined. First, since each corresponding feature results in two measurements, z^r and z^l, and each measurement contains two elements for the u and v pixel co-ordinates respectively, a single feature actually provides 4 independent measurements. As such, the full measurement vector is constructed from feature point set Y is

$$y_Y = [z^{lu}_{d_1} \; z^{lv}_{d_1} \; z^{ru}_{d_1} \; z^{rv}_{d_1} \; \ldots \; z^{lu}_{d_{N_Y}} \; z^{lv}_{d_{N_Y}} \; z^{ru}_{d_{N_Y}} \; z^{rv}_{d_{N_Y}}]^T \in \mathbb{R}^k$$

where $k = 4N_Y$ is the number of independent measurements provided by the feature point set Y. In a similar fashion, the nonlinear measurement model constructed from set Y, $h(x)_Y : \mathbb{R}^n \mapsto \mathbb{R}^k$ is given as

$$h(x)_Y = [\Theta_x(d_i)^l \; \Theta_x(d_i)^r \; \ldots \; \Theta_x(d_{N_Y})^l \; \Theta_x(d_{N_Y})^r]^T$$

Preprint submitted to 2014 IEEE International Conference on Robotics and Automation. Received September 15, 2013.

B. Outlier Rejection

Loosely speaking, an outlier can be defined as a measurement which, by some measure, is inconsistent with the majority of the measurements. In a typical RANSAC approach, a hypothesis model is constructed from a sample set of measurements randomly selected from the full measurement set. Each measurement is then compared to the hypothesis model and is added to an inlier consensus set if an application specific fitness measure is satisfied. In the case of visual odometry, the reprojection error for the feature must be less than a given threshold to be condsidered an inlier. The entire process of constructing hypothesis models from random samples and determining the inliers is repeated for a fixed number of iterations, while keeping track of the largest consensus set. A final model for the system can then be generated using the largest inlier consensus set.

III. Parity Space Fault Detection

In this section, a brief overview of the parity space fault detection methodology is presented. For a thorough discussion on parity space fault detection and isolation, the reader is directed to [23], [24]. Although the approach is demonstrated using a linear measurement model, it should be noted that a nonlinear measurement model, $h(x)$, can be linearized about an operating point, x_0, to generate an approximate linear model,

$$\hat{y} = \bar{H}x$$

where $\bar{H}$ is the linearized measurement model, $\bar{H} = \frac{\partial h}{\partial x}|_{x_0}$, and $\hat{y} = y - h(x_0)$ are the shifted measurements according to the operating point. Although the linearization introduces a state dependent measurement matrix to the parity space approach, the linearization error is negligible assuming the chosen operating point is sufficiently close to the true state [23]. In the case of egomotion estimation using visual odometry, performing the linearization assuming zero movement between successive frames is acceptable, provided the framerate of the camera is sufficiently high.

A. Parity Vector Generation

Suppose that k measurements are observed at any given time-step. For a system with n states, the measurement equation is defined as

$$y = Hx + e + f \tag{2}$$

where $y \in \mathbb{R}^k$ is the measurement vector, $H \in \mathbb{R}^{k\times n}$ is the measurement model, $x \in \mathbb{R}^n$ is the state vector, $e \in \mathbb{R}^k$ is the additive measurement noise vector, and $f \in \mathbb{R}^k$ is the fault vector which models the error in the measurements under the assumption that the measurement set contains an outlier. If y contains no outliers, the fault vector f is the zero vector. The additive measurement noise, e, is assumed to be drawn from a Gaussian distribution with zero mean and measurement noise covariance $Q \in \mathbb{R}^{k\times k}$, or $e \sim \mathcal{N}(0, Q)$. Assuming the measurements are independent and the noise magnitude between measurements is similar, Q can be simply modelled as

$$Q = \sigma_e I_{k\times k}$$

where $\sigma_e \in \mathbb{R}$ is the measurement noise variance.

The parity space approach seeks to transform the measurements into a vector space, known as the *parity space* where outlier rejection can be performed. The projection of the measurements into the parity space is a linear transformation defined by the matrix $V \in \mathbb{R}^{(k-n)\times k}$, and the column space of V is defined as the parity space. For the parity space approach, it is required that V is an orthogonal matrix that is also orthogonal to H, or

$$VH = 0 \tag{3}$$

$$VV^T = I_{(k-n)} \tag{4}$$

Although any V which satisfies Equations 3 and 4 is sufficient, a simple method for computing V exists. Suppose the matrix $W \in \mathbb{R}^{(k)\times k}$ is given as

$$W = I - H(H^TH)^{-1}H^T$$

If W is post multiplied by H, it is clear that Equation 3 is satisfied. To satisfy Equation 4, a Gram-Schmidt orthogonalization procedure is performed on W, resulting in an orthogonal matrix V. The *parity vector*, $p \in \mathbb{R}^{(k-n)}$, is defined as the projection of the measurements, y, onto the parity space, and can be used to determine if the measurement vector contains an outlier. Once V is known, the parity vector is calculated as

$$p = Vy \tag{5}$$

If the expression for the measurements from 2 is substituted into 5, the parity vector is

$$p = VHx + Ve + Vf$$

Due to Equation 3, the term VHx is zero, and the state is removed from the problem. The resulting parity vector is

$$p = Ve + Vf$$

With the noise assumption of $E[e] = 0$ and $E[ee^T] = Q$, if there are no outliers in the data, then the expected value and covariance of the parity vector is

$$E[p] = 0$$

$$\begin{aligned} Cov[p] &= E[pp^T] - E[p]E[p^T] \\ &= E[Vee^TV^T] \\ &= VQV^T \end{aligned}$$

Thus, when no outliers are present in the data, the parity vector is normally distributed, or $p \sim \mathcal{N}(0, VQV^T)$. In order to model an inconsistency in the measurements, assume that the i^{th} measurement is an outlier. The outlier can be modelled using the vector f by setting the i^{th} component of f, f_i to a non-zero value. Then, the parity vector is distributed as

$$p \sim \mathcal{N}(v_i f_i, VQV^T)$$

Preprint submitted to 2014 IEEE International Conference on Robotics and Automation. Received September 15, 2013.

where v_i is the i^{th} column of V. Therefore, when an outlier is present the mean of the parity vector is shifted by magnitude f_i, in the fault direction given by v_i.

B. Outlier Detection Test

Under the assumption of no outliers in the measurement vector, the magnitude of the parity vector is generally small as defined by the measurement noise Q. Conversely, when an outlier is present, the magnitude of the parity vector is dominated by the size of the fault component, f_i. Thus, the magnitude of the parity vector can be used to construct an outlier detection test statistic, and is given as

$$\lambda = p^T p$$

The test statistic, λ is chi-squared (χ^2) distributed with $k-n$ degrees of freedom, therefore, a χ^2 test can be used to determine the presence of an outlier. Suppose a false alarm probability of α_{fa} is desired. The critical threshold, δ, which satisfies the probability $P(\lambda > \delta) = \alpha_{fa}$, can be determined using a χ^2 distribution look up table. Next, define the null hypothesis, $\mathcal{H}_0$, and the alternative hypothesis, $\mathcal{H}_1$ as

$$\mathcal{H}_0 : \ y \text{ contains no outliers}$$
$$\mathcal{H}_1 : \ y \text{ contains an outlier}$$

If $\lambda \leq \delta$, the measurements are consistent, there are no outliers present and the null hypothesis is accepted. Else, if $\lambda > \delta$, the test indicates that a there is an inconsistency in the measurements due to an outlier and the null hypothesis is rejected.

C. Probability of Missed Detection

Define the probability of missed detection, P_m as the probability that the outlier detection test declares no outliers are present when, in fact, the measurement vector contains an outlier. Generally, P_m is difficult to determine accurately as it requires the integration of the parity vector distribution which has been shifted by the fault vector $f_i v_i$, over the hypersphere defined by the detection test threshold, δ. However, an upper bound on P_m can be determined by marginalizing the parity vector distribution in the direction of the fault vector, resulting in a one dimensional distribution which can be easily integrated. The upper bound on P_m is given as

$$P_m < \int_{-\delta}^{\delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{\rho - f_i}{\sigma_e^2})} d\rho \tag{6}$$

Although this is an acceptable upper bound, in cases where the parity space is relatively low ($(k-n) = 2$), it is possible to compute a tighter approximation for the upper bound by further integrating Equation 6 over the disk defined by the detection threshold, δ.

IV. Proposed Approaches

In addition to the detection of an outlier within the measurement vector, the *isolation* of the inconsistent measurement is also possible within the parity space framework [23], [24]. However, the approach is only well suited for the isolation of one outlier, where as in most visual odometry applications, the removal of multiple outliers is necessary. As such, performing the parity space outlier isolation analysis is unsuitable, and a different approach is required to handle outliers from camera feature matches.

A. Parity Group Outlier Rejection

To remove multiple outliers from the measurement vector, the Parity Group Outlier Rejection (GPOR) strategy is proposed. A similar approach is discussed in [18], however the method of selection for the group is unclear. The GPOR strategy simply divides the measurement vector into groups of size g, and performs the parity space outlier detection test on each individual group. If the test fails, all features within the group are discarded. The selection of the group size is an important parameter, as a large group size will lead to a high rate of false positives and discard many good measurements, while a group size that is too small will increase the likelihood of including a group that contains multiple outliers that are in agreement in the final measurement vector. Excessive discarding of good measurements will also occur if the outlier ratio is large, and so GPOR can be thought of as conservative in its assessment of outliers. The method is therefore most applicable for data with reasonably low outlier ratios and allows for rapid outlier removal that scales linearly with the number of measurements.

B. Parity Initialized RANSAC

Since the GPOR algorithm is best suited to measurements with a lower outlier ratio, another approach, such as RANSAC, is required for situations where a large percentage of the observations are corrupt. Although promising, it is evident that the performance of the RANSAC algorithm is heavily dependent on the inlier to outlier ratio of the measurement data. Suppose the inlier ratio for the data is given by τ and the number of random samples used to generate the hypothesis model is c. Then, after l RANSAC iterations, the probability of the algorithm returning an incorrect model, P_b, is given by

$$P_b = (1 - \tau^c)^l$$

Figure 1 illustrates how P_b is affected by the expected inlier ratio of the data for $c = 3$, which is the minimum number of 3D feature correspondences required to estimate the motion between two consecutive stereo-pair frames. It is clear that when the outlier percentage is high, many RANSAC iterations are required in order to generate a correct model with sufficiently high certainty.

In certain applications, running the RANSAC algorithm for many iterations becomes computationally expensive. In the case of the presented visual odometry algorithm, a nonlinear optimization must be performed at every iteration in order to generate the candidate model from the random sample set. Furthermore, determination of the inliers using the pixel reprojection error is on the order of $O(k)$. As the number of measurements, k, can become quite large in practice, performing fewer RANSAC iterations is desirable to improve computational performance.

Preprint submitted to 2014 IEEE International Conference on Robotics and Automation. Received September 15, 2013.

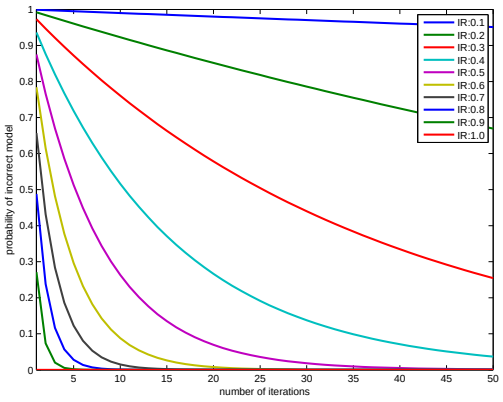

Fig. 1. Probability of RANSAC generating an incorrect model for $c = 3$. Probabilities are computed for inlier ratios ranging from 10% to 100% (IR:0.1 to IR:1.0).

In order to reduce the number of required RANSAC iterations without sacrifice of solution quality, we propose the Parity Initialised RANSAC (PI-RANSAC) algorithm, where a parity space consistency check on the randomly selected measurement set is performed at each iteration. Denote the candidate sample set as S, where S is selected randomly from a uniform distribution over the entire 3D feature set, Y, such that $S \sim \mathcal{U}(Y)$ and $|S| = N_S$. Recall that each 3D point in S, d_i, is taken from the previous image-pair, and has an associated feature correspondence, z_{d_i}, in the current image-pair. To that end, define y_S as the measurement vector containing only the feature correspondences according to the points from S, or

$$y_S = [z^{lu}_{d_1}\ z^{lv}_{d_1}\ z^{ru}_{d_1}\ z^{rv}_{d_1}\ \dots\ z^{lu}_{d_{N_S}}\ z^{lv}_{d_{N_S}}\ z^{ru}_{d_{N_S}}\ z^{rv}_{d_{N_S}}]^T$$

Instead of using y_S to calculate the model, as one would with a typical RANSAC approach, the PI-RANSAC algorithm performs a parity space consistency check to determine if y_S contains an outlier. If the consistency check fails, another random sample set is selected. Once the consistency test is passed, the model is computed using y_S, and the inlier set for the current iteration, Y^i_I, is determined. A feature is considered an inlier if the pixel reprojection error is less than a predefined threshold, ϵ. By using the parity based check on the random measurement sample, the computational burden of model computation and determination of the full inlier set at each iteration is replaced with a simple parity space test that is on the order of $O(c^3)$. It is important to note that although the parity check scales cubically, mainly due to computation of the V matrix, in typical applications $c << k$. As such, the benefit of the PI-RANSAC algorithm is more pronounced as the number of measurements increase. The entire process is repeated for l_p iterations, where upon termination, the algorithm returns the largest inlier set, Y_{best}. The full PI-RANSAC approach as applied to the visual odometry problem is summarized in Algorithm 1.

The effectiveness of the PI-RANSAC approach is dependent on the accuracy of the parity space test of the random sample set. If the parity check was 100% effective, meaning $P_{md} = 0$, only one PI-RANSAC iteration would be required. However, such an approach is only possible if the distribution of the fault vector is completely known, which is an unreasonable assumption in practice. In the case of outlier rejection for visual odometry, the magnitude of the fault, f_i, is unknown. Using Equation 6, the probability of missed detection as a function of the normalized magnitude of the fault vector is displayed in Figure 2. To generate the figure, the test detection threshold is set to $\delta = 6\sigma_e$, and the fault vector magnitude is varied from $f_i = 0$ to $f_i = 10\sigma_e$. It is evident that the larger the difference between the detection threshold and the expected value of the fault magnitude, the lower the probability of a missed detection will be.

Using the probability of a missed detection, it is possible to determine the probability the PI-RANSAC algorithm will return a bad model as a function of the number of iterations. After performing l_p iterations, the probability of an incorrect model is

$$P_b = (P^{f_i}_m)^{l_p}$$

where $P^{f_i}_m$ denotes the probability of a missed detection from the parity space test, given a fault vector magnitude of f_i. Figure 3 illustrates the probability of returning an incorrect model as a function of the PI-RANSAC iterations performed,

Algorithm 1 Given the full set of 3D features from the previous image pair, Y, and an initial state estimate x_0, perform outlier rejection and return the best inlier set of features, Y_{best}, using the PI-RANSAC method.

$Y_{best} \leftarrow \emptyset$
for $i = 1 : l_p$ **do**
  $Y^i_I \leftarrow \emptyset$
  perform parity check:
  while $\lambda > \delta$ **do**
    $S \sim \mathcal{U}(Y)$
    $y_S \leftarrow [z^{lu}_{d_1}\ z^{lv}_{d_1}\ z^{ru}_{d_1}\ z^{rv}_{d_1} \dots z^{lu}_{d_{N_S}}\ z^{lv}_{d_{N_S}}\ z^{ru}_{d_{N_S}}\ z^{rv}_{d_{N_S}}]^T$
    $\bar{H} \leftarrow \frac{\partial}{\partial x} h(x)_S|_{x_0}$
    $\hat{y} \leftarrow y_S - h(x_0)_S$
    $W \leftarrow I - \bar{H}(\bar{H}^T\bar{H})^{-1}\bar{H}^T$
    $V \leftarrow \text{gramSchmidt}(W)$
    $p \leftarrow V\hat{y}$
    $\lambda \leftarrow p^T p$
  end while
  determine solution using sample set S:
  $x^* = \underset{x\in\mathbb{R}^n}{\text{argmin}}\ \Lambda(x)_S$
  create inlier set for current iteration:
  for all $d_i \in Y$ **do**
    if $\|z^l_{d_i} - \Theta_{x^*}(d_i)^l\|^2 + \|z^r_{d_i} - \Theta_{x^*}(d_i)^r\|^2 < \epsilon$ **then**
      $Y^i_I \leftarrow Y^i_I \cup d_i$
    end if
  end for
  store largest inlier set:
  if $|Y^i_I| > |Y_{best}|$ **then**
    $Y_{best} \leftarrow Y^i_I$
  end if
end for

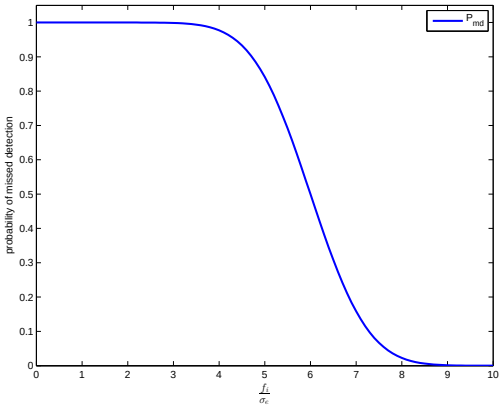

Fig. 2. Probability of a missed detection using the parity space test for a detection threshold of $\delta = 6\sigma_e$. Probabilities are computed as a function of the expected value of the fault vector magnitude, ranging from $f_i = 0$ to $f_i = 10\sigma_e$.

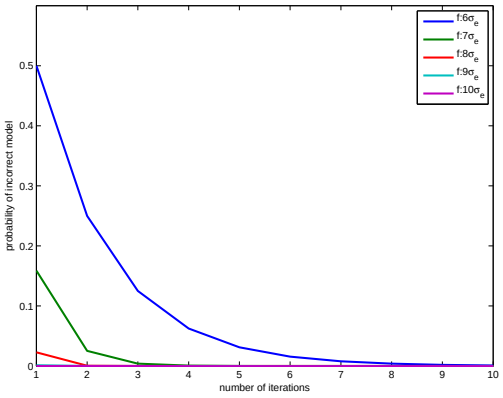

Fig. 3. Probability of the PI-RANSAC algorithm returning an incorrect model as a function of the number of performed iterations with a detection threshold of $\delta = 6\sigma_e$. Plots are shown for normalized fault vector magnitudes ranging from $f_i = \delta$ to $f_i = 10\sigma_e$.

for a range of fault vector magnitudes. It can be seen that even when the expected value of the fault vector magnitude is equal to the parity space test detection threshold ($f_i = \delta$), the probability of returning an incorrect model is 0.0009 after only 10 iterations.

V. Experimental Results

In order to validate the proposed methods, a series of experiments are performed using camera data from the KITTI Vision Benchmark Dataset [22]. The image data is collected using two Point Grey Flea 2 cameras mounted on an automotive platform driving in urban environments. Ground truth of the vehicle motion is recorded by an OXTS RT 3003 integrated GPS/IMU, capable of providing position data at an accuracy of approximately 10 cm. The outlier rejection methods are applied to the LIBVISO2 implementation of the visual odometry method outlined in [8]. The mapping between the naming convention of the data sequences to the names as provided from the KITTI Dataset are shown in Table I.

The experiments consist of performing the visual odometry algorithm while applying different outlier rejection techniques. The Group Party Outlier Rejection (GPOR) technique, as well as the PI-RANSAC algorithm performed for 10 iterations (PR-10), is compared to the RANSAC algorithm performed with 10, 50, 100 and 1000 iterations (denoted in the tables as R-10,R-50,R-100,R-1000). The incremental motion computed between frames using each outlier rejection techniques is integrated over the test sequence to recover the egomotion of the camera, and is compared to the provided ground truth. To determine the effect which the number of features has on the outlier rejection techniques, the experiments are performed using small and large feature sets for the test sequences. The number of features is controlled by increasing the maximum number of allowable features maintained by the bucketing algorithm [7]. Other than the number of features and RANSAC iterations, all other parameters for the visual odometry are set to the default values. For the PI-RANSAC implementation, the detection threshold is set to $\delta = 0.5$ and for the GPOR method the group size was set to $g = 3$. The results from the experiments for the small and large feature sets are presented in Tables II and III, respectively.

Sequence Name	KITTI Dataset
S01	2011_09_26_drive_0067
S02	2011_09_30_drive_0016
S03	2011_09_30_drive_0020
S04	2011_09_30_drive_0027
S05	2011_09_30_drive_0033

TABLE I
Sequence Name Mapping

For the small feature number experiment, the PI-RANSAC approach computes the solution with the least average position error, for all five test sequences. Although intuitively it seems that the average position error should decrease as the number of RANSAC iterations is increased, in some cases a large number of RANSAC iterations returns an *accidental* largest consensus set which does not accurately reflect the state of the system [25]. To that end, the PI-RANSAC approach is compared to the *best* solution discovered by RANSAC over 10, 50, 100 and 1000 iterations. In Tables II and III, the percent improvement of the PI-RANSAC solution over the best RANSAC solution is computed using the bold quantities. It is evident that the PI-RANSAC approach is very effective, as after only 10 iterations it was able to compute a solution with an average 44.45% increase in accuracy and an average 68.95% improvement in run-time compared to the best RANSAC solutions.

A similar trend is seen with the experiments performed on the large feature data set, with the exception of R-1000 for the S02, sequence, where the average position error for R-1000 is marginally better compared to PR-10. Significant computational benefit is seen for the PI-RANSAC method on the large feature sets, as in the best case a 97.19% speed-up was observed when compared to RANSAC. Although the GPOR method did not produce the solutions with the best accuracy relative to ground truth, the average position errors are indeed comparable to both PI-RANSAC and RANSAC,

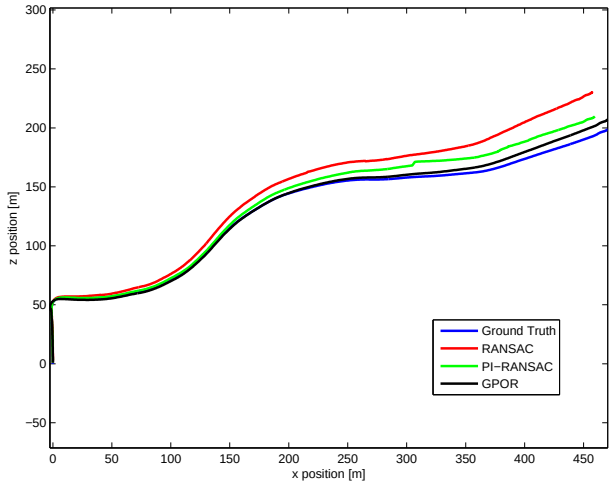

Fig. 4. Vehicle motion generated using visual odometry and outlier rejection techniques for S01.

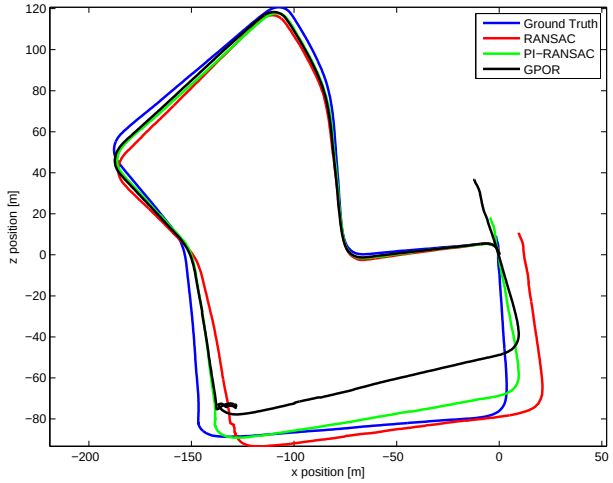

Fig. 5. Vehicle motion generated using visual odometry and outlier rejection techniques for S04.

and consistently achieved the best run-time in all test cases. Example plots comparing the calculated vehicle motion trajectory using GPOR, R-10, and PR-10 for sequences S01, S04, S05 are presented in Figures 4, 5 and 6, respectively.

One issue with evaluation of the integrated visual odometry motion is that a poor solution in the middle of the sequence of position estimates will affect all subsequent positions. Thus, the relative error between the frame to frame positions and relative motion derived from the ground truth are computed, and summary statistics are presented in Table V for both small and large feature counts. As the grouping for the GPOR method is based purely on the measurement vector, in situations where the outlier ratio is high, the test may have a tendency to reject a high percentage of good measurements which results in a low quality solution. However, it is apparent that for the sequences with large feature counts, the GPOR method produced the solutions with the smallest average relative error. The result suggests that the test sequences with large numbers of features mitigate the

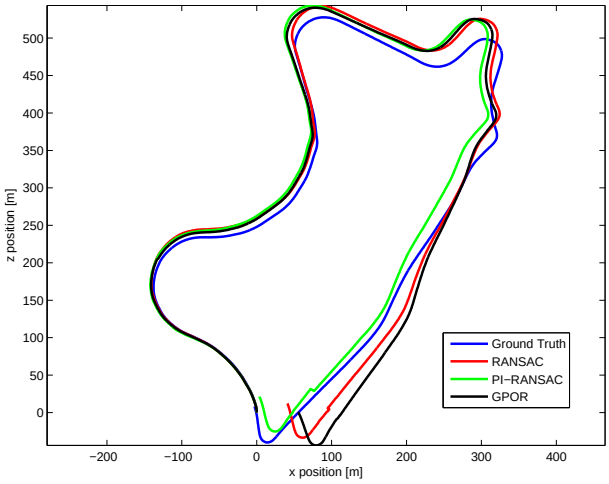

Fig. 6. Vehicle motion generated using visual odometry and outlier rejection techniques for S05.

effects of feature deprivation that occur in the smaller feature sets, and lead to more consistent position estimates over the entire trajectory. As a result, for methods where large feature sets are being used and where the outlier ratio is low, a simple test such as the GPOR method can rapidly produce accurate solutions, and can outperform RANSAC methods in both accuracy and runtime.

	S01	S02	S03	S04	S05
avg. feature count	299	211	234	292	245
Average Position Error (m)					
GPOR	8.59	2.64	8.82	13.75	22.00
PR-10	**7.39**	**0.89**	**4.39**	**3.43**	**19.81**
R-10	20.19	3.39	8.78	9.75	28.75
R-50	14.36	2.03	9.53	**9.05**	27.63
R-100	**12.14**	3.17	7.95	10.04	**26.04**
R-1000	13.46	**1.90**	**7.83**	12.89	30.49
Improvement (%)	39.13	53.16	43.93	62.10	23.92
Average Run Time (ms)					
GPOR	0.95	0.70	0.74	0.85	0.79
PR-10	**2.80**	**4.00**	**4.80**	**4.00**	**4.60**
R-10	1.30	1.00	1.00	1.30	1.00
R-50	6.40	4.80	5.10	**5.60**	5.00
R-100	**12.10**	90.0	9.50	11.00	**13.00**
R-1000	68.70	**30.0**	**40.0**	70.60	56.00
Speed-Up (%)	76.86	86.67	88.00	28.57	64.62

TABLE II

ACCURACY AND RUNTIME COMPARISON USING A SMALL FEATURE SET (LESS THAN 300 FEATURES ON AVERAGE).

VI. CONCLUSION

Performing outlier rejection within the parity space framework is beneficial as it allows for a computationally inexpensive detection of inconsistent measurements without explicit computation of the system state. This work presents two outlier rejection approaches, the GPOR and PI-RANSAC method that are based within the parity space framework, and applies them to a stereo camera based visual odometry application. The GPOR algorithm allows for rapid outlier rejection by performing parity space tests to subgroups of

	S01	S02	S03	S04	S05
avg. feature count	2070	1320	1231	1594	1408
Average Position Error (m)					
GPOR	3.18	3.09	6.24	12.23	22.75
PR-10	**6.61**	**1.62**	**6.82**	**6.20**	**18.67**
R-10	14.31	3.56	14.66	10.72	23.29
R-50	9.75	3.01	8.79	10.01	**21.00**
R-100	10.97	1.57	**7.75**	10.03	21.79
R-1000	**9.26**	**1.56**	10.70	**7.95**	21.60
Improvement (%)	28.62	-3.85	12.00	22.02	11.11
Average Run Time (ms)					
GPOR	6.50	4.10	3.70	5.10	4.50
PR-10	**9.90**	**8.80**	**8.30**	**12.10**	**9.30**
R-10	8.20	5.40	5.00	6.30	5.60
R-50	38.10	25.10	22.80	30.90	**26.40**
R-100	70.90	45.30	**43.40**	58.40	50.90
R-1000	**352.60**	**188.60**	161.60	**341.30**	256.30
Speed-Up (%)	97.19	95.33	80.88	96.45	64.77

TABLE III

ACCURACY AND RUNTIME COMPARISON USING A LARGE FEATURE SET (GREATER THAN 1000 FEATURES ON AVERAGE).

Average Relative Position Error (m)					
	S01	S02	S03	S04	S05
avg. feature count	299	211	234	292	245
GPOR	**0.049**	0.090	0.084	0.060	**0.082**
PR-10	0.051	0.086	**0.060**	**0.042**	0.088
R-10	0.091	0.103	0.097	0.061	0.115
R-50	0.064	0.094	0.096	0.065	0.104
R-100	0.060	0.091	0.079	0.060	0.104
R-1000	0.601	**0.085**	0.078	0.072	0.103
Average Relative Position Error (m)					
avg. feature count	2070	1320	1231	1594	1408
GPOR	**0.038**	**0.083**	**0.064**	**0.056**	**0.082**
PR-10	0.051	0.096	0.087	0.058	0.083
R-10	0.069	0.111	0.078	0.061	0.089
R-50	0.052	0.091	0.094	0.064	0.102
R-100	0.052	0.073	0.083	0.061	0.117
R-1000	0.048	0.087	0.101	0.060	0.107

TABLE IV

COMPARISON OF AVERAGE FRAME BY FRAME RELATIVE ERROR FOR BOTH SMALL AND LARGE AVERAGE FEATURE COUNTS.

features of the measurement vector, resulting in a fast outlier removal strategy that scales linearly with the number of measurements. However, since the performance of the GPOR algorithm degrades in data with larger outlier ratios, the PI-RANSAC algorithm is proposed, which improves upon RANSAC by performing a fast parity space consistency test on the random sample of measurements selected when performing RANSAC. Since the parity test ensures only consistent random samples are evaluated for the maximum inlier set, the PI-RANSAC algorithm is able to provide an accurate solution in fewer iterations when compared to RANSAC. The proposed methods are implemented on the LIBVISO2 visual odometry library and are successfully validated using stereo image data from the KITTI Vision dataset. Future work includes more thorough analysis of parity space approach with nonlinear models and application of the proposed outlier rejection methods on problems such as SLAM and point cloud registration.

REFERENCES

[1] P. Huber, *Robust Statistics*. New York: Wiley, 1974.

[2] P. J. Rousseeuw, "Least median of squares regression," *Journal of the American statistical association*, vol. 79, no. 388, pp. 871–880, 1984.

[3] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, Jan. 1972.

[4] C. H. Tong and T. D. Barfoot, "Batch heterogeneous outlier rejection for feature-poor slam," in *International Conference on Robotics and Automation (ICRA)*. Shanghai, China: IEEE, May 2011, pp. 2630–2637.

[5] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[6] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, ISBN: 0521540518, 2004.

[7] B. Kitt, A. Geiger, and H. Lategahn, "Visual odometry based on stereo image sequences with ransac-based outlier rejection scheme," in *Intelligent Vehicles Symposium (IV)*, San Diego, CA, June 2010, pp. 486–492.

[8] A. Geiger, J. Ziegler, and C. Stiller, "Stereoscan: Dense 3d reconstruction in real-time," in *Intelligent Vehicles Symposium (IV)*. Baden-Baden, Germany: IEEE, June 2011, pp. 963–968.

[9] B. Morisset, R. B. Rusu, A. Sundaresan, K. Hauser, M. Agrawal, J.-C. Latombe, and M. Beetz, "Leaving flatland: Toward real-time 3d navigation," in *International Conference on Robotics and Automation (ICRA)*. Kobe,Japan: IEEE, may 2009, pp. 3786 –3793.

[10] G. Klein and D. Murray, "Parallel tracking and mapping for small AR workspaces," in *International Symposium on Mixed and Augmented Reality (ISMAR)*. Nara, Japan: IEEE, November 2007, pp. 225–234.

[11] P. H. S. Torr and A. Zisserman, "MLESAC: a new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138 – 156, 2000.

[12] O. Chum, J. Matas, and S. Obdrzalek, "Enhancing RANSAC by generalized model optimization," in *Asian Conference on Computer Vision (ACCV)*, Jeju, Korea, January 2004, pp. 812–817.

[13] J. Matas and O. Chum, "Randomized ransac with sequential probability ratio test," in *International Conference on Computer Vision (ICCV)*. Beijing, China: IEEE, October 2005, pp. 1727–1732.

[14] O. Chum and J. Matas, "Randomized RANSAC with T(d,d) test," in *British Machine Vision Conference (BMVC)*, Cardiff, UK, September 2002, pp. 448–457.

[15] I. E. Potter and M. C. Sunman, "Threshold-less redundancy management with arrays of skewed instruments," AGARD, Tech. Rep. AGARDOGRAPH-224 (pp 15-25), 1977.

[16] M. D. Asok Ray and J. Deyst, "Fault detection and isolation in a nuclear reactor," *Journal of Energy*, vol. 7, no. 1, pp. 79–85, 1983.

[17] S. Hall, P. Motyka, E. Gai, and J. J. Deyst, "In-flight parity vector compensation for FDI," *Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 5, pp. 668–676, 1983.

[18] D. Tornqvist, T. Schon, and F. Gustafsson, "Detecting spurious features using parity space," in *International Conference on Control Automation, Robotics and Vision (ICARCV)*. Hanoi, Vietnam: IEEE, December 2008, pp. 353–358.

[19] D. Nistr, O. Naroditsky, and J. Bergen, "Visual odometry for ground vehicle applications," *Journal of Field Robotics*, vol. 23, no. 1, pp. 3–20, 2006.

[20] J.-P. Tardif, Y. Pavlidis, and K. Daniilidis, "Monocular visual odometry in urban environments using an omnidirectional camera," in *International Conference on Intelligent Robots and Systems (IROS)*. Nice, France: IEEE, September 2008, pp. 2531–2538.

[21] A. Comport, E. Malis, and P. Rives, "Accurate quadrifocal tracking for robust 3d visual odometry," in *International Conference on Robotics and Automation (IROS)*, San Diego,CA, October 2007, pp. 40–45.

[22] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[23] F. Gustafsson, "Statistical signal processing approaches to fault detection," *Annual Reviews in Control*, vol. 31, no. 1, pp. 41–54, 2007.

[24] F. Gustafsson and F. Gustafsson, *Adaptive filtering and change detection*. Wiley Londres, 2000, vol. 1.

[25] Q. Fan, "Matching slides to presentation videos," Ph.D. dissertation, Department of Computer Science, University of Arizona, 2008.

Preprint submitted to 2014 IEEE International Conference on Robotics and Automation. Received September 15, 2013.