2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

RANSAC Matching: Simultaneous Registration and Segmentation

Shao-Wen Yang, Chieh-Chih Wang and Chun-Hua Chang

Abstract— **The iterative closest points (ICP) algorithm is widely used for ego-motion estimation in robotics, but subject to bias in the presence of outliers. We propose a random sample consensus (RANSAC) based algorithm to simultaneously achieving robust and realtime ego-motion estimation, and multi-scale segmentation in environments with rapid changes. Instead of directly sampling on measurements, RANSAC matching investigates initial estimates at the *object* level of abstraction for systematic sampling and computational efficiency. A soft segmentation method using a multi-scale representation is exploited to eliminate segmentation errors. By explicitly taking into account the various noise sources degrading the effectiveness of geometric alignment: sensor noise, dynamic objects and data association uncertainty, the uncertainty of a relative pose estimate is calculated under a theoretical investigation of scoring in the RANSAC paradigm. The improved segmentation can also be used as the basis for higher level scene understanding. The effectiveness of our approach is demonstrated qualitatively and quantitatively through extensive experiments using real data.**

I. Introduction

Ego-motion estimation in dynamic environments is one of the most fundamental problems in mobile robotics, which is the problem of determining the pose of a robot relative to its previous location. It is not easily achievable as there are two motions involved: the motions of moving objects and the motion of the robot itself. A large body of work in computer vision over the last decade has been concerned with the extraction of ego-motion information from image sequence [1–3]. The performance of ego-motion estimation depends on the consistency between observations at successive time steps, and can be degraded in the presence of outliers. The motivation of this work is intended to provide a robust realtime solution to the problem of relative pose estimation in non-rigid scenes.

The iterative closest points (ICP) algorithm [4, 5], which is based on least squares minimization, has been widely used for aligning range images. However, conventional least squares approaches are subject to bias in the presence of outliers. Since the introduction of the ICP algorithm, many variants have been proposed upon the basic ICP concept. Rusinkiewicz and Levoy [6] proposed a combination of ICP variants optimized for high speed with a point-to-plane error metric. Pfister et al. [7] introduced a weighted matching algorithm to estimate the transformation by matching successive range scans. Pfister and Burdick [8] described a multi-scale point and line-based representation of range scans to improve efficiency of scan matching. Minquez et al. [9] used a new metric distance in the robot's configuration space. Bosse and Zlot [10] presented an iterative scan matching technique using an extended Kalman filter maintaining all the vehicle poses. More recently, we proposed a RANSAC-based segment matching approach [11] to estimate a robot's ego-motion in dynamic environments. However, conventional approaches cannot deal with dynamic objects. Segment-based approaches are subject to imperfect segmentation.

As the key development in robotics has been the adoption of probabilistic approaches, many recent state-of-the-art robotic systems employ probabilistic techniques for robotic perception. However, in developing a practical relative pose estimation algorithm, it can be difficult to quantify the effectiveness of a relative pose estimate. Gutmann and Schlegel [12] described a comparison on the covariances from several scan matching approaches in indoor environments. These approaches work fine in office-like environments, particularly with orthogonal or rectilinear walls, but are infeasible for unstructured environments. Lu and Milios [13] have shown how a covariance matrix for the ICP algorithm can be estimated directly from the corresponding pair of points. Unfortunately, the uncertainty estimates are too conservative and do not correspond to reality. Bengtsson and Baerveldt [14] presented to calculate the covariance matrix by estimating the Hessian matrix of the error function minimized by the iterative dual correspondence (IDC) algorithm. The approximation mandates predefined constant offsets in translation and rotation for evaluating the minimization error. Wang and Thorpe [15] proposed a hierarchical object based representation for simultaneous localization and mapping (SLAM) which estimates the uncertainty using a sampling-based approach. The registration process is performed repetitively with random initial estimates. Censi [16] proposed a Laplace approximation to calculate the covariance matrix for scan matching algorithms.

In the computer vision literature, random sample consensus (RANSAC) [17] is one of the most effective algorithms for model fitting to data containing a significant percentage of gross errors. It is an iterative method to estimate parameters of a mathematical model from a set of observed data which

This work was supported in part by the Taiwan National Science Council under Grant 96-2628-E-002-251-MY3, Grant 96-2218-E-002-035, Grant 97-2218-E-002-017, and Grant 98-2218-E-002-006, in part by the Excellent Research Projects of the National Taiwan University, in part by Taiwan Micro-Star International, in part by Compal Communications, Inc., and in part by Intel.

Shao-Wen Yang is with the Department of Computer Science and Information Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan any@pal.csie.ntu.edu.tw

Chieh-Chih Wang is with Faculty of the Department of Computer Science and Information Engineering, and the Graduate Institute of Networking and Multimedia, National Taiwan University bobwang@ntu.edu.tw

Chun-Hua Chang is with the Department of Computer Science and Information Engineering, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei, 10617 Taiwan alan@pal.csie.ntu.edu.tw

contains outliers. A theoretical investigation of scoring under a simple inlier-outlier model is performed to discriminate outliers from the inlier model. The RANSAC paradigm which is capable of fitting data containing a significant percentage of gross errors is advanced in its effectiveness and efficiency, and particularly applicable to scene analysis. RANSAC-based approaches in the computer vision literature mainly focus on segmenting a range image by feature extraction and parameter fitting [18, 19]. Chen et al. [20] applies a rigidity constraint for feature search in a data set. A random-selection strategy is repeated until a solution is found. These approaches can fail to reliably compute a robot's relative pose, particularly when registering range scans in unknown environments, which are rather impoverished in the localized landmarks [5].

Based on the RANSAC-based segment matching approach [11], the paper introduces the RANSAC matching algorithm utilizing a multi-scale representation of range images. To provide robustness against poor segmentation and moving objects, RANSAC matching solves the problem of relative pose estimation at the object level of abstraction, in which data association uncertainty and segmentation uncertainty are tackled simultaneously. The uncertainty of a relative pose estimate is calculated under a theoretical investigation of scoring in the RANSAC paradigm. The multi-scale segmentation can also supply a significant preprocessing step for many robotics applications, such as modeling of non-stationary environments [21] and moving entity tracking from a moving vehicle [22]. RANSAC matching does not employ any geometric features which are often environment dependent. It also inherits the computational efficiency and probabilistic robustness from the RANSAC paradigm. The proposed approach is implemented and tested in the context of ego-motion estimation and SLAM using laser range data. Experimental results demonstrate the effectiveness and robustness of our algorithm.

II. Random Sample Consensus

First of all, we review the foundation and probabilistic formulation of RANSAC. Classical techniques for parameter estimation optimize the fit of a functional description to all of the presented data. The RANSAC procedure is opposite to that of conventional smoothing technique. Rather than using as much of the data as possible to obtain an initial solution and then attempting to eliminate the invalid data points, RANSAC uses as small an initial data set as feasible and enlarges this set with consistent data when possible [17].

RANSAC uses the geometric distribution in statistics which models the discrete distribution: the probability distribution of the number X of Bernoulli trials needed to get one success, supported on natural numbers $\mathbb{N}$. If the probability of success on each trial is b, then the probability that the k-th trial is the first success is

$$\begin{aligned} \Pr(X = k) &= (1-b)^{k-1} b && (1) \\ &= (1-w^n)^{k-1} w^n && (2) \end{aligned}$$

where w is the probability that any selected data point is within the error tolerance of the model, and n is the number of good data points required to determine the model, for all $k \in \mathbb{N}$. To ensure with probability p that at least one of the random selections is an error-free set of n data points, we must expect to make k selections, where

$$\begin{aligned} (1-b)^k &\leq (1-p), && (3) \\ k &\geq \log(1-p)/\log(1-b). && (4) \end{aligned}$$

RANSAC is effective for model fitting, particularly when a significant percentage of data are outliers. It is ideally suited for applications in range image analysis. A detailed derivation and a comprehensive description can be found in Fischler and Bolles [17]. The RANSAC formulation contains two remaining unspecific parameters n and w which are highly relevant to characteristics of data. In the next section, we will derive the parameters for the ego-motion estimation problem where non-static objects can be rejected as outliers.

III. RANSAC Segment Matching

Ego-motion estimation can be performed using range image registration algorithms in the computer vision literature. To ensure against the possibility of the final consensus set being compatible with an incorrect model, the size of data points per selection should be greater than or equal to three for determining the pose transformation, including translation and rotation. However, one of the most difficult problems in range image processing is data association. Every single measurement is featureless. Researchers usually apply the closest point association rule to register data points with unknown data association, such as the ICP algorithm [5].

A. Segmentation

In the RANSAC paradigm, a number of random samples consisting of small sets are taken from an observation. A first attempt is to generate random samples directly from all measurements of an observation. Closest point association often yields good estimate for data containing a mass of points. However, registration performs very poorly on data containing few points and often results in ambiguity. Sampling directly on all measurements also requires comparatively large size of data points to preserve sufficient shape information for registration. For example, if $w = 0.5$ and $n = 5$, then $b = 0.03125$. To obtain a 99% assurance of making at least one error-free selection, by Equation 4 we have $k \geq 146$. It is time-consuming and computationally intractable for realtime applications, even though five points are still insufficient for obtaining a good registration result.

Instead of direct sampling on measurements, we propose to use a higher level data representation – *segment* – to achieve reliable registration and realtime performance. An observation is segmented and further split into sets of measurements representing objects. Specifically, objects are extracted by segmenting a range image into densely sampled parts. Here, we use a distance criterion to segment measurements into objects.

B. Matching

In the classical RANSAC paradigm, letting o be a feature and h be some hypothesis, the effectiveness of each (o, h) is examined and represented using a binary score. Specifically, if (o, h) is an inlier pair, the score s_h of the hypothesis h is incremented. As segments might be of significantly different sizes, a binary score is insufficient to describe the quantity of an association between two segments. Let z be an observation and z^i be the i-th segment in z. Comparing to the classical RANSAC process, the score s_h^i of each segment z^i is supported on $\mathbb{N}$ and the effectiveness of the pair (z^i, h) is represented by a natural number.

1) Sampling: To build consensus sets, z is segmented and represented as a collection of segments $z = \cup_i z^i$. First, segments are randomly selected with probabilities proportional to their sizes. A hypothesis h is generated by matching the selected n segments with the reference model $\bar{z}$.

2) Scoring: To obtain the score s_h^i of a segment z^i, the effectiveness of (z^i, h) is examined by checking if (y, h) is an inlier pair for all $y \in z^i$. The score s_h^i of a segment z^i is defined as the number of measurements $y \in z^i$ which are located within neighborhoods of measurements in the reference model $\bar{z}$. Here, (y, h) is judged as an inlier pair if and only if the measurement y transformed to the global coordinate by the hypothesis h is located within a neighborhood d of some measurement in the reference model $\bar{z}$. Specifically, the score s_h^i is incremented if the pair (y, h) is judged as an inlier pair. Therefore, we have

$$s_h = \sum_{\{i|z^i \subseteq z\}} s_h^i \tag{5}$$

$$= \sum_{z^i \subseteq z} \sum_{y \in z^i} \mathbf{1}_h(y) \tag{6}$$

where $\mathbf{1}_h(y)$ is an indicator function indicating whether or not (y, h) is an inlier pair. When the process finishes, the hypothesis with the highest score is output as the best transformation ψ. In this work, d is 0.3 meter.

The parameter n should be carefully determined and take into account the tradeoff between efficiency and reliability, and the characteristics of the data. For matching segments with the reference model, one segment is usually sufficient to preserve the shape information of the environment unless an environment is composed of line segments which result in ambiguity. It is clear that the higher the value n, the higher the probability at least one hypotheses is an inlier, and thus the reliability increases. Letting $n = 2$ and $w = 0.5$, according to Equation 4, to obtain a 99% assurance of making at least one error-free selection, the number k of selections must be greater than or equal to 17, which is computationally sufficient for realtime applications.

However, the present segmentation approach can fail as the characteristics of an environment are subject to change over time. Objects may be mis-merged with a high threshold, while using a too low threshold results in over-segmentation. Figure 1 gives an example illustrating the segmentation issue and Figure 2(c) shows the visual image. On the top of these figures, a static object and a moving object are very close to each other. In Figures 1(b), 1(c) and 1(d), the two objects are mis-segmented and merged together as they are not spaced far enough apart from each other, whereas, in Figures 1(e), 1(f) and 1(g), the two objects are properly segmented but most of the objects are split into fragmented segments.

IV. Multi-scale RANSAC Matching

In this section, we describe the proposed RANSAC matching algorithm in which segmentation issues are resolved using a multi-scale representation. As segmentation is used as a preprocessing step in segment-based approaches, the segmentation errors introduced by the hard decisions bring difficulties to the segment matching algorithm. A scale tree representing *objects* of varying sizes in a range image is proposed to eliminate the segmentation errors.

A. Multi-scale Representation

We define a scale tree comprising collections of segments extracted from a range image at multiple scales. Edges inheriting parent-child relationships on the tree are established for segments at different scales for which the child segment at a finer scale is subsumed in the parent segment at a coarser scale. A scale tree is constructed in a top-down manner. The process is started by extracting segments at the coarsest scale and then iteratively extracting finer segments by splitting from the coarser segments. Let $z^{j,i}$ denote the i-th segment extracted at the j-th scale. The vertex set of a scale tree can then be defined as $\cup_j \cup_i z^{j,i}$. Figure 1 illustrates an example of the multi-scale representation of a range image. In this work, the segmentation thresholds are 12, 9, 6, 3, 1.5, and 0.75 meter respectively.

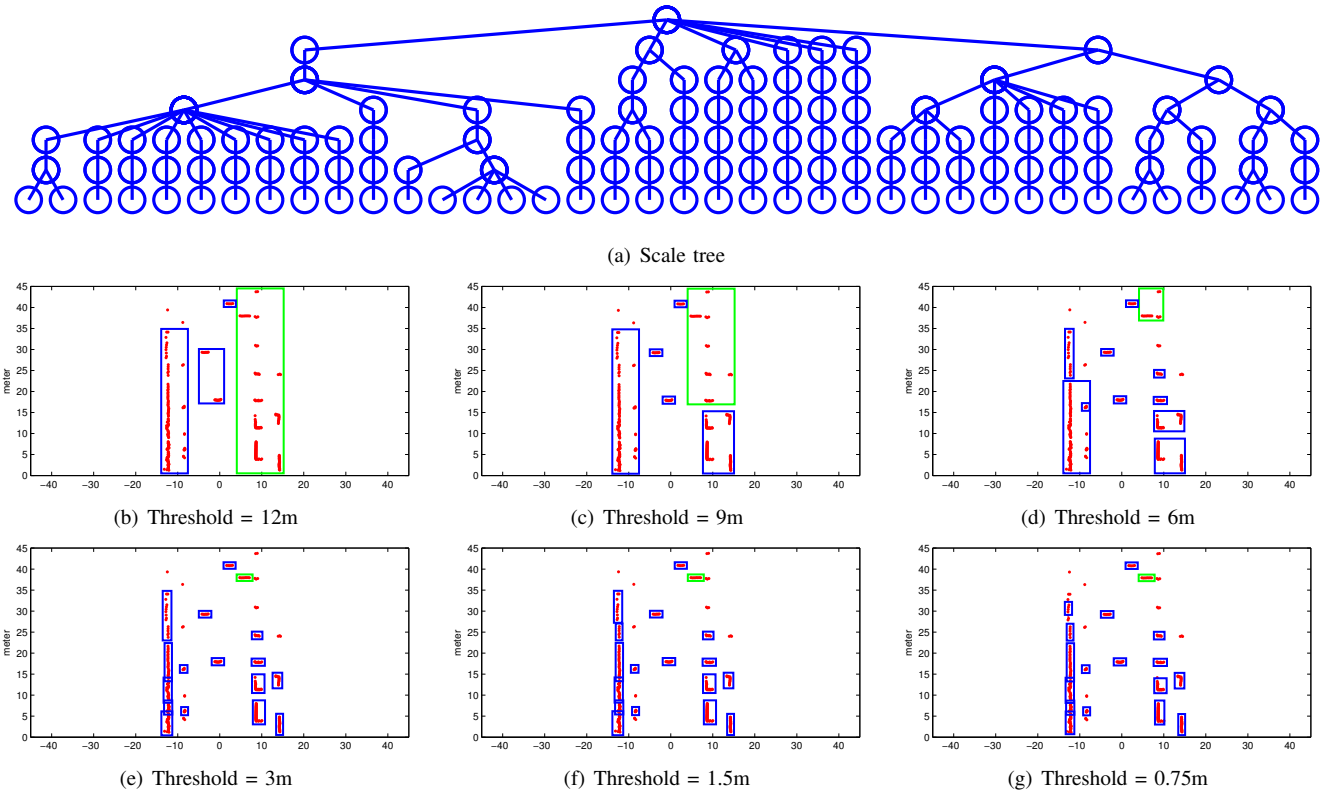
B. Multi-scale Matching

To avoid under-segmentation and over-segmentation, we take advantage of the multi-scale representation wherein an observation z is represented as collections of segments extracted at multiple scales. Instead of making hard decisions in segmentation, we propose a soft segmentation method using the multi-scale representation to simultaneously deal with data association uncertainty and segmentation uncertainty. All segments in a scale tree are taken into account in the sampling stage of the RANSAC process. Comparing to the segment matching approach, the consensus set is generated from randomly selected n segments from all segments within the scale tree.

Constructed in this manner, the uncertainty in segmentation is modeled directly in the process of building consensus sets. It is assumed that the probability of a measurement belonging to a segment at some scale is uniformly distributed, which can be expressed as

$$p(z^{j,i}|y) \sim \mathcal{U}(\ell), \quad \forall z^{j,i} \ni y \tag{7}$$

where $\mathcal{U}(\ell)$ denotes the uniform distribution on $\mathbb{N}_1^\ell$, and ℓ is the number of scales used for constructing the scale tree.



(a) Scale tree

(b) Threshold = 12m

(c) Threshold = 9m

(d) Threshold = 6m

(e) Threshold = 3m

(f) Threshold = 1.5m

(g) Threshold = 0.75m

Fig. 1. Multi-scale representation. In (a), the scale tree of the range image is depicted. The segmentation result at each scale is visualized in (b), (c), (d), (e), (f) and (g), in which the observation is shown in red dots, the segments are shown in rectangles, and segments with less than four points are not depicted for clarity. The green rectangle indicates that a static object and a moving object are mis-merged in (b), (c) and (d), and properly segmented in (e), (f) and (g).

A hypothesis is then generated by aligning the sampled segments with the reference model $\bar{z}$. To obtain the score s_h of a hypothesis h, the effectiveness of (z, h) is examined by accumulating the effectiveness of each measurement in the observation z. Thus, we can rewrite Equation 6 in a more general form as

$$s_h = \sum_{y \in z} \mathbf{1}_h(y) \tag{8}$$

where $\mathbf{1}_h(y)$ is an indicator function indicating whether or not (y, h) is an inlier pair. The same as previously, the RANSAC process then outputs the hypothesis with the highest score as the ego-motion estimate ψ.

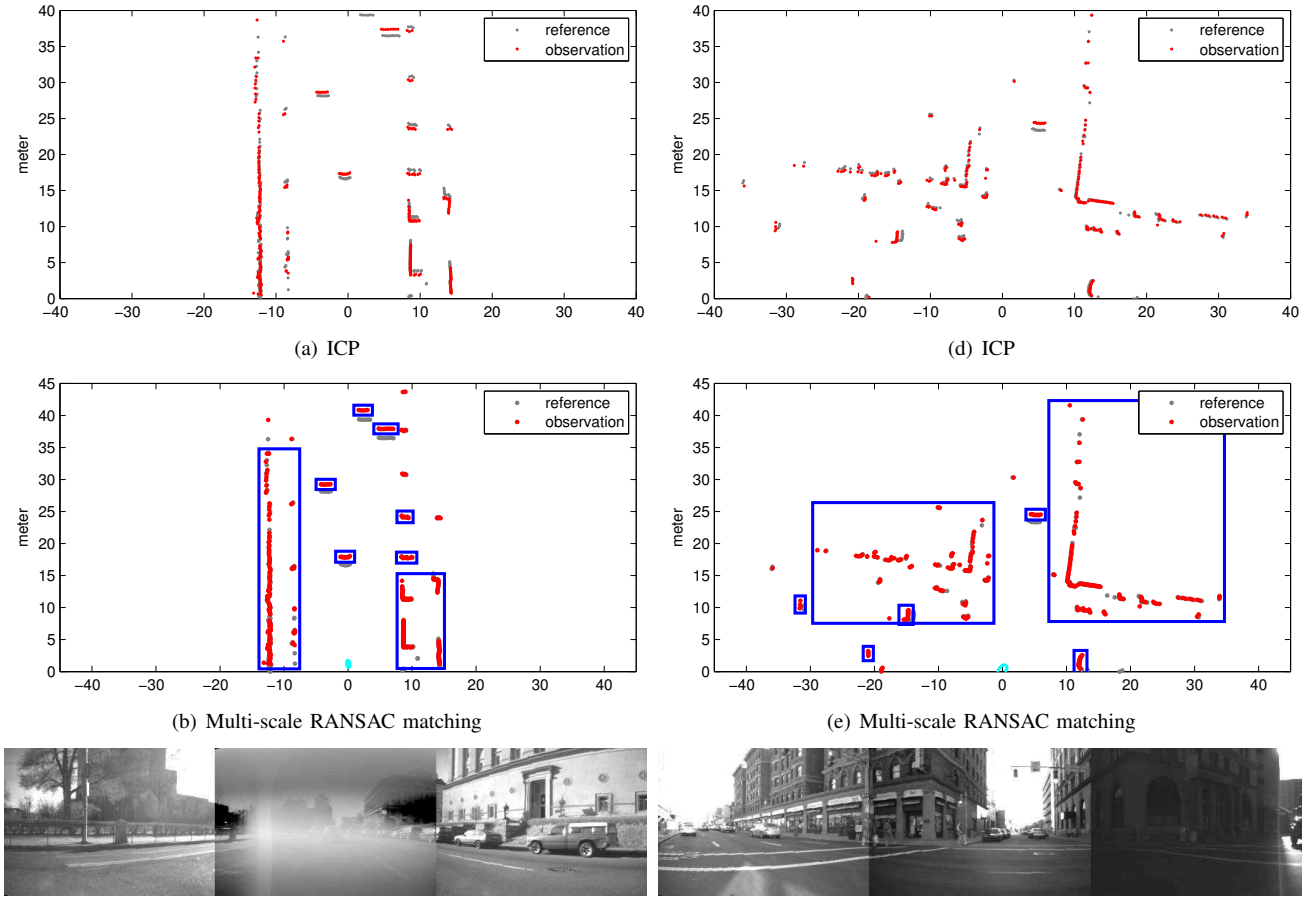
In the RANSAC process, the multi-scale representation is utilized for hypothesis generation. In addition to the assumption that at least 50% of the measurements in an observation are static, to cope with the segmentation issue, we further take into account the segmentation uncertainty which is assumed uniformly distributed. We define that a segment is consistent if and only if its measurements undergo the same motion. Let v be the probability a measurement is consistently segmented. Observe the segmentation results given in Figure 1 that the objects on the top which are mis-merged in Figures 1(b), 1(c) and 1(d), and consistently segmented in Figures 1(e), 1(f) and 1(g). As a result, for each measurement of an observation, there exists some scale such that scales below yield a consistent segment whereas thresholds above do not. Hence, according to Equation 7, we assume $v = 0.5$, without loss of generality. By Equation 4, to obtain a 99% assurance of making at least one error-free selection, we have

$$\begin{aligned} k &\geq \log(1-p)/\log(1-b) \\ &= \log(1-p)/\log(1-(w \cdot v)^n) \\ &= \log(1-0.99)/\log(1-0.0625) \\ &\geq 72 \end{aligned} \tag{9}$$

where the probability of success on each trial $b = (w \cdot v)^n$ takes into account the segmentation uncertainty, as segments at multiple scales are employed in the RANSAC process.

C. Multi-scale Segmentation

To achieve better segmentation, we further perform edge deletion on the scale tree to remove inconsistent segments. The tree is split in a bottom-up manner. Each segment $z^{j,i}$ in the observation z is transformed into the global coordinate with the transformation ψ. We define $\omega(z^{j,i})$ as the effectiveness of the association between the segment $z^{j,i}$ and the reference model $\bar{z}$, which is evaluated by calculating the percentage of measurements in $z^{j,i}$ within neighborhoods of measurements in $\bar{z}$. Specifically, if a segment is static, it is probably associated with some segment in the reference model in a relative great proportion, unless it be moving or



(a) ICP (d) ICP

(b) Multi-scale RANSAC matching (e) Multi-scale RANSAC matching

(c) Visual Image (f) Visual Image

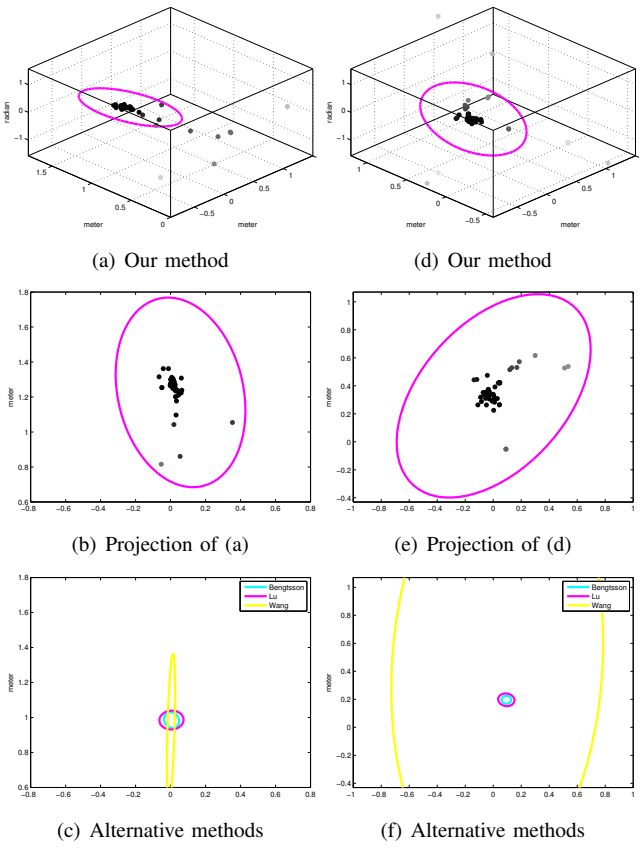
Fig. 2. Multi-scale RANSAC matching. The ICP results are shown in (a) and (d) while the corresponding RANSAC matching results are shown in (b) and (e). Visual images for these two examples are depicted in (c) and (f), respectively. In (a), (b), (d) and (e), the observation is shown in red dots and the reference model is shown in gray dots. In (b) and (e), the segments are shown in blue rectangles, and the uncertainty ellipse at around the origin is shown in cyan. Regarding the segmentation results only the root vertices of scale trees are depicted for clarity.

occluded. Thus, if $\omega(z^{j,i})$ is less than some proportion ϕ, the vertex corresponding to $\omega(z^{j,i})$ and its descendants are split from the scale tree and forms a new tree. The split operation is performed from the leaf vertices of a scale tree, and cuts all edges connecting to the split vertex from the scale tree. In our implementation, the value of ϕ is 50%, which is the only parameter has to be chosen in this work, in addition to RANSAC parameters.

Figure 2 demonstrates the effectiveness of the proposed RANSAC matching algorithm and the multi-scale segmentation. Figures 2(a) and 2(d) show the ICP results, Figures 2(b) and 2(e) depict the RANSAC matching results, and Figures 2(c) and 2(c) present the corresponding visual images. The most pressing issue of the conventional scan matching algorithms is that they do not explicitly cope with outliers. It can be seen that the RANSAC matching algorithm outperforms the ICP algorithm. The ICP algorithm can possibly converge to local minima in the presence of moving objects, particularly in dynamic scenes. This is typical for least squares approaches in which the quadratic penalty allows a outlier far apart from the true solution to bias the final result [17]. Specifically, the ICP algorithm uses the whole observation to obtain a solution of the relative pose, whereas the RANSAC matching algorithm samples as few static objects within a range image as feasible, and evaluates hypotheses by checking the consistency between the sampled objects and the reference model. The multi-scale segmentation can also supply a significant preprocessing step for a variety of robotics applications, and be used as the basis for higher level scene understanding.

V. Uncertainty Estimation

Covariance estimation is necessary to quantify the uncertainty of a relative pose estimate. Most probabilistic processes, such as extended Kalman filters (EKFs) and particle filters (PFs), should be aware of the level of confidence in the state estimates. A realistic covariance estimate is also necessary for further combining the relative pose estimates with additional odometric or inertial measurements [23]. For example, in a Kalman filter framework, the contribution of measurements from different sensors to the state estimate is weighted by the Kalman gains whose values depend on the covariances of all the sources of information contributing to the filter.



(a) Our method (d) Our method

(b) Projection of (a) (e) Projection of (d)

(c) Alternative methods (f) Alternative methods

Fig. 3. Uncertainty estimation. The uncertainty estimates of the examples given in Figures 2(b) and 2(e) are visualized in (a) and (d), respectively. The projections of (a) and (d) onto the x-y coordinates are shown in (b) and (e). The ellipse shows 99% confidence bound of the covariance estimate and the RANSAC hypotheses are shown in gray dots. The uncertainty estimates from other methods are shown in (c) and (f).

Let H be the set of hypotheses generated by the RANSAC process. We calculate the weight w_h for each RANSAC hypothesis $h \in H$ by normalizing the scores of all hypotheses, which can be expressed as

$$w_h = \frac{s_h}{\sum_{h \in H} s_h} \tag{10}$$

such that $w_h \in [0, 1]$ and $\sum_{h \in H} w_h = 1$. For clarity, let $h^\star \in H$ denote the hypothesis selected by the RANSAC process which yields the highest score. The covariance matrix C is estimated by measuring the statistical dispersion about the hypothesis $h^\star$ with the highest score. Specifically, statistical dispersion is variability or spread in a variable or a probability distribution in statistics. Each entry $C_{i,j}$ of a covariance C on row i and column j can thus be calculated as

$$C_{i,j} = \frac{\sum_{h \in H} w_h \left(\psi_{h,i} - \psi_{h^\star,i}\right)\left(\psi_{h,j} - \psi_{h^\star,j}\right)}{1 - \sum_{h \in H} w_h^2} \tag{11}$$

where ψ_h is the transformation generated by hypothesis h, and $\psi_{h,i}$ is the i-th element of ψ_h.

Figures 2(b) and 2(e) also visualize the mean and the covariance estimates. Enlargements of the uncertainty ellipses are shown in Figures 3(a) and 3(b), and Figures 3(d) and 3(e), respectively, in which the RANSAC hypotheses are also depicted. It can be seen that while moving forward along a straight road, as illustrated in Figures 2(b), 3(a) and 3(b), the estimated relative pose is more uncertain in the forward direction. In Figures 2(e), 3(d) and 3(e), while making a left turn, a strong correlation is evident between the forward and sideward directions.

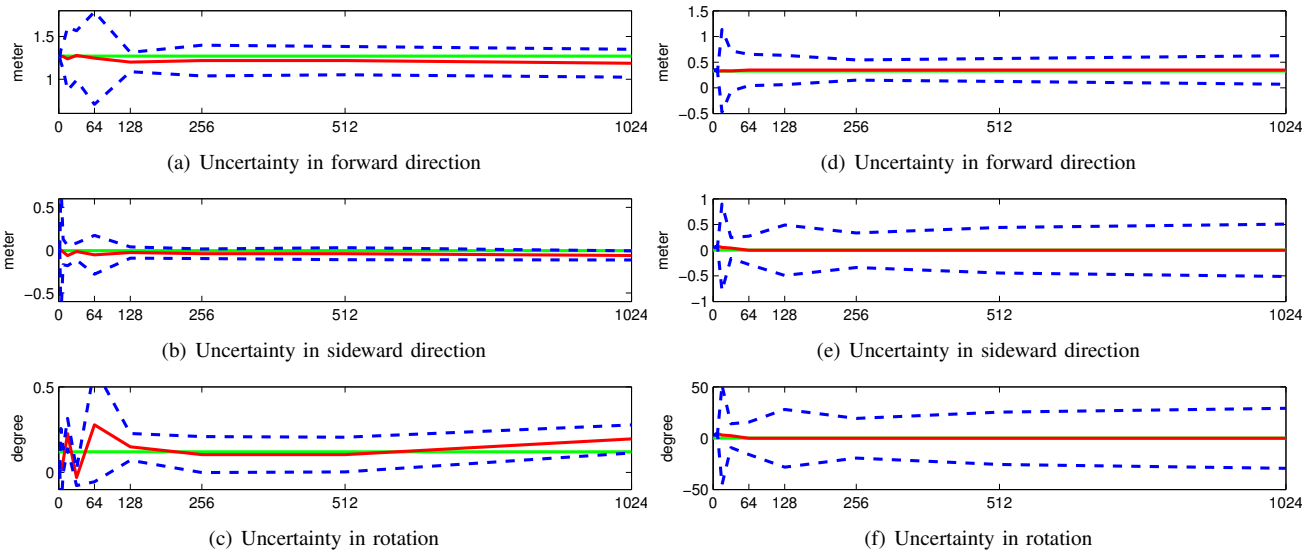
VI. Experimental Results

The RANSAC matching algorithm was evaluated extensively using real range data [24]. The travel distance of the data set is approximately 5 kilometers. We compare the relative pose estimates from the ICP algorithm and the RANSAC matching algorithm with the ground truth, which is provided from the onboard inertial measurement system of the vehicle [15]. Though either the odometric or inertial data are actually not the ground truth, they can still provide sufficiently locally accurate information of the vehicle's motion to some degree within a short period of time. In our implementation, the data are processed at around 7.5 Hz. Table I shows the performance improvements of RANSAC matching. The ICP algorithm used is similar to the high-speed variant introduced by Rusinkiewicz and Levoy [6] with a point-to-plane error metric, constant-weighting, a distance threshold for rejecting pairs, and the standard select-match-minimize ICP iterations. As can be seen, the improvement of segment-based matching is subject to imperfect segmentation. The use of the multi-scale representation is particularly effective to simultaneously deal with objects of significantly sizes. The soft segmentation method is robust to errors in segmentation at any particular resolution. Our approach yields a 20% improvement in translation and a 8% improvement in rotation.

TABLE I
Improvement over ICP

Threshold	12m	9m	6m	3m	1.5m	0.75m	multi
Translation	6%	9%	15%	12%	13%	13%	20%
Rotation	-11%	-11%	0%	8%	8%	8%	8%

Furthermore, we show the effectiveness and convergency of RANSAC matching. In Figure 3, the uncertainty estimates using different methods are shown. Figures 3(a), 3(b), 3(d), and 3(e) show the uncertainty estimates using our method, and Figures 3(c) and 3(f) show the uncertainty estimates using alternative methods [13–15]. It can be seen that the uncertainty estimates from Bengtsson and Baerveldt's and Lu and Milios' methods are too optimistic and do not correspond to reality, in which covariance matrices are estimated from the residuals of the corresponding points. Wang and Thorpe's method can provide reasonable estimates as it takes into account the uncertainty in data association, which perform scan matching repetitively using random initial estimates. Figure 4 demonstrates the effectiveness and convergency of RANSAC matching. The RANSAC matching algorithm is performed on the same data with varying number of samples to estimate the distribution of the relative pose estimate.



(a) Uncertainty in forward direction
(b) Uncertainty in sideward direction
(c) Uncertainty in rotation
(d) Uncertainty in forward direction
(e) Uncertainty in sideward direction
(f) Uncertainty in rotation

Fig. 4. Effectiveness and convergency. The x-axis denotes the number of samples used in the RANSAC process. The red line and the blue dashed line show the mean and the 99% confidence bound, respectively, of a relative pose estimate. The green line shows the ground truth obtained from the inertial measurement system. The uncertainty estimate of the example given in Figure 2(b) is shown in (a), (b) and (c), and the uncertainty estimate of the example given in Figure 2(e) is shown in (d), (e) and (f). Note that the vertical scales of (c) and (f) are considerably different.

In both of the examples, there are about 100 segments at multiple scales in total. Enumerating all combinations of segments at multiple scales is computationally infeasible for realtime applications. It can be seen that the significant property of RANSAC matching is the trade-off between increased representation power and computational overhead. Initially, the accuracy of the method increases with the number of samples. As increased number of samples, the estimate tends to converge to a specific steady state within a neighborhood of the ground truth. The property ensures the convergency and efficiency of our algorithm. Figure 5 shows an empirical analysis of accuracy and convergency for RANSAC matching. We assumed that an estimate is accurate if it differs by at most 0.1 meter in forward and sideward directions, and 1 degree in rotation from the ground truth, and an uncertainty estimate is converged if it differs by at most 0.1 meter in forward and sideward directions, and 1 degree in rotation from the estimate obtained from using 1024 samples. The ample experimental results in consequence are consistent with the theoretical derivation of multi-scale RANSAC matching given in Equation 9. Along with the processing rate, the feasibility of RANSAC matching for realtime estimation is also confirmed.

In summary, the experiments indicate that our algorithm provides robust relative pose estimates in terms of ego-motion estimation and uncertainty estimation. In comparison with the alternative approaches [13–15], our algorithm yields more accurate relative pose estimates and also provides more consistent uncertainty estimates.

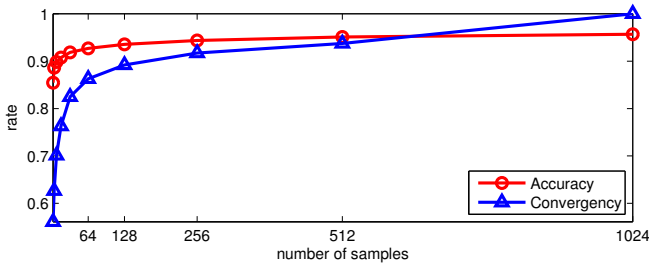


Fig. 5. Success rate. The x-axis denotes the number of samples used in the RANSAC process.

VII. Conclusion And Future Work

We introduced a novel approach to solve the problem of registration and segmentation of range images. RANSAC matching solves the problem at the object level of abstraction by exploiting a multi-scale representation to model objects of varying sizes. The main contribution of this paper is to propose a robust realtime algorithm which takes into account data association uncertainty and segmentation uncertainty simultaneously in the RANSAC paradigm. Instead of applying random initial estimates, the RANSAC process can investigate initial estimates systematically and score hypotheses statistically. By representing a range image as segments at multiple scales, our approach overcomes a range of limitations possessed by least squares approaches, such as the inability to consistently estimate the uncertainty of a relative pose estimate, and poor degradation to outliers. RANSAC matching does not employ any geometric features which are often environment dependent. It also inherits the computational efficiency and probabilistic robustness from the RANSAC paradigm. The feasibility and effectiveness of the proposed approach have been demonstrated using real data collected in urban environments without incorporating odometry. Experimental results show that our approach outperforms alternative approaches.

In the future, we plan to integrate spatial and temporal

information into a unified theoretical framework to deal with the occlusion problem. Furthermore, theoretically, the proposed approach can be extended to the matching problem for 3D point clouds without much difficulty. Practically, the computation time perhaps can be far from realtime due to the explosion of data points and the multi-scale tree nodes. We are investigating possibilities to integrate additional information from a 3D laser range finder or a stereo camera, such as intensity and color. Future investigation will also include applying the framework of RANSAC matching to a broader range of problems.

REFERENCES

[1] O. D. Faugeras, F. Lustman, and G. Toscani, "Motion and structure from motion from point and line matches," in *Proceedings of the IEEE International Conference on Computer Vision*, London, UK, 1987, p. 25.

[2] D. J. Heeger and A. D. Jepson, "Subspace Methods for Recovering Rigid Motion I: Algorithm and implementation," *International Journal of Computer Vision*, vol. 7, no. 2, pp. 95–117, 1992.

[3] K. J. H. James R. Bergen, P. Anandan and R. Hingorani, "Hierarchical Model-Based Motion Estimation," in *Proceedings of the European Conference on Computer Vision*, vol. 588. Santa Margherita Ligure, Italy: Springer, 1992, pp. 237–252.

[4] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 2, pp. 239–256, 1992.

[5] F. Lu and E. Milios, "Robot Pose Estimation in Unknown Environments by Matching 2D Range Scans," *Journal of Intelligent and Robotic Systems*, vol. 18, no. 3, pp. 935–938, 1997.

[6] S. Rusinkiewicz and M. Levoy, "Efficient Variants of the ICP Algorithm," in *3rd International Conference on 3D Digital Imaging and Modeling*, Quebec City, Quebec, Canada, May 2001, pp. 145–152.

[7] S. T. Pfister, K. L. Kriechbaum, S. I. Roumeliotis, and J. W. Burdick, "Weighted Range Sensor Matching Algorithms for Mobile Robot Displacement Estimation," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Washington, DC, May 2002, pp. 1667–1674.

[8] S. T. Pfister and J. W. Burdick, "Multi-scale Point and Line Range Data Algorithms for Mapping and Localization," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Orlando, Florida, May 2006.

[9] J. Minguez, L. Montesano, and F. Lamiraux, "Metric-Based Iterative Closest Point Scan Matching for Sensor Displacement Estimation," *IEEE Transactions on Robotics*, vol. 22, no. 5, pp. 1047–1054, October 2006.

[10] M. Bosse and R. Zlot, "Map Matching and Data Association for Large-Scale Two-dimensional Laser Scan-based SLAM," *The International Journal of Robotics Research*, vol. 27, no. 6, pp. 667–691, June 2008.

[11] S.-W. Yang and C.-C. Wang, "Multiple-Model RANSAC for Ego-Motion Estimation in Highly Dynamic Environments," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Kobe, Japan, May 2009, pp. 3531–3538.

[12] J.-S. Gutmann and C. Schlegel, "AMOS: Comparison of Scan Matching Approaches for Self-Localization in Indoor Environments," in *Proceedings of the First Euromicro Workshop on Advanced Mobile Robots*, Kaiserslautern, Germany, October 1996, pp. 61–67.

[13] F. Lu and E. Milios, "Globally Consistent Range Scan Alignment for Environment Mapping," *Autonomous Robots*, vol. 4, no. 4, pp. 333–349, 1997.

[14] O. Bengtsson and A.-J. Baerveldt, "Localization in Changing Environments – Estimation of a Covariance Matrix for the IDC Algorithm," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Maui, HI, USA, October 2001, pp. 1931–1937.

[15] C.-C. Wang and C. Thorpe, "A Hierarchical Object based Representation for Simultaneous Localization and Mapping," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, September 2004.

[16] A. Censi, "An accurate closed-form estimate of ICP's covariance," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Roma, Italy, April 2007.

[17] 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, 1981.

[18] X. Yu, T. D. Bui, and A. Krzyzak, "Robust Estimation for Range Image Segmentation and Reconstruction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 5, pp. 530–538, 1994.

[19] H. Wang and D. Suter, "MDPE: A Very Robust Estimator for Model Fitting and Range Image Segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 139–166, September 2004.

[20] C.-S. Chen, Y.-P. Hung, and J.-B. Cheng, "RANSAC-based DARCES: A New Approach to Fast Automatic Registration of Partially Overlapping Range Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 11, pp. 1229–1234, 1999.

[21] D. Anguelov, R. Biswas, D. Koller, B. Limketkai, and S. Thrun, "Learning Hierarchical Object Maps Of Non-Stationary Environments With Mobile Robots," in *Proceedings of the Conference on Uncertainty in Artificial Intelligence*. San Francisco, CA: Morgan Kaufmann, 2002, pp. 10–17.

[22] C.-C. Wang, C. Thorpe, and S. Thrun, "Online Simultaneous Localization And Mapping with Detection And Tracking of Moving Objects: Theory and Results from a Ground Vehicle in Crowded Urban Areas," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Taipei, Taiwan, September 2003.

[23] S. I. Roumeliotis and J. W. Burdick, "Stochastic Cloning: A generalized framework for processing relative state measurements," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Washington, DC, May 2002.

[24] C.-C. Wang, D. Duggins, J. Gowdy, J. Kozar, R. MacLachlan, C. Mertz, A. Suppe, and C. Thorpe, "Navlab SLAMMOT Datasets," http://www.csie.ntu.edu.tw/ bobwang/datasets.html, May 2004, Carnegie Mellon University.