

Adaptive Scale Robust Segmentation for 2D Laser Scanner

Ruben Martinez-Cantin, José A. Castellanos, Juan D. Tardós and J. M. M. Montiel

*Dept. Informática e Ingeniería de Sistemas
Instituto de Investigación en Ingeniería de Aragón (I3A)
Universidad de Zaragoza
María de Luna 1, E-50018 Zaragoza, Spain
Email: {rmcantin, jacaste, tardos, josemari}@unizar.es*

Abstract—This paper presents a robust algorithm for segmentation and line detection in 2D range scans. The described method exploits the multimodal probability density function of the residual error. It is capable of segmenting the range data in clusters, estimate the straight segments parameters, and estimate the scale of inliers error noise successfully, despite of high level of spurious data. No prior knowledge about the sensor and object properties is given to the algorithm. The mode seeking is based on mean shift algorithm, which has been widely used and tested in 3D laser scan segmentation, machine learning and pattern recognition applications. We show the reliability of the technique with experimental indoor and outdoor manmade environment. Compared with classical methods, a good compromise between false positive, false negative, wrong segment split and wrong segment merge is achieved, with improved accuracy in the estimated parameters.

I. INTRODUCTION

Robot navigation, localization and mapping requires to deal with a high amount of data, sometimes redundant and commonly noisy. In manmade environments the assumption of flat surfaces is highly reliable (walls, doors, tables, bookshelves, etc.). In those cases, the typical platform is a mobile robot with a built-in 2D rangefinder sensor moving on the floor. Then, flat surfaces from real world are transformed into straight lines on the sensor plane.

Feature extraction in robotics takes advantage of human capability of abstraction which reports three main advantages: (i) *data compression*: features use a simple parametric model to represents the high amount of data, (ii) *denoising*: since the model is estimated from several measurements, the uncertainty and bias of final estimation is lower than any of the measurements, (iii) *distinguishability*: high level features and complex structures are easier to identify and match (despite partial observation).

Real surfaces are not planar but have relief and hence it is an issue to define when a cluster of points, not exactly over a line, can be approximated as straight segment. A key concept in the approximation is the scale, defined by the noise standard deviation. It is worth noting that this scale is a property of each scene segment, and hence it has to be determined from the data corresponding to the segment. For example, in Figure 1, the cluster corresponding to the ivy covered wall on the left has a bigger scale than the clusters corresponding to the concrete walls on the right, despite being

sensed approximately at the same relative location with respect to the sensor, this shows how scale, is not only determined by the sensor but also by the scene. Right bottom part of the figure shows a successful clustering and the corresponding σ estimate (magnified 30 times). Classical benchmark methods [1] do not care about scale computation. Uncertainty models were constructed based on sensor noise, considering a perfect scene model, and considering the noise scale as an input to the algorithm.

Thereinafter, without losing generality, observations follows the next hypothesis based on the structural assumptions

- Data points \mathcal{Y} satisfy $\mathcal{Y} \in \mathbb{R}^2$, because we suppose a robot moving on flat surface (floor), with a 2D rangefinder parallel to the floor.
- Environment structures are planes perpendicular to the floor, that appears as straight lines in the working planes.
- Noise distribution v is perpendicular to the estimated lines $v_i \sim p(v|\mathcal{Y}, \theta)$ following a gaussian distribution $v_i \sim \mathcal{N}(0, \sigma)$.
- The standard deviation σ of the error distribution corresponds to the scale, that is, the data dispersion of inliers; which can vary from one segment to another.

Once the scan has been segmented in clusters corresponding to the straight segments, the line fitting with these hypothesis can be optimally solved using Total Least Squares [2]. The problem is that clustering and scale estimation are coupled and have influence on the line fitting, being a unique spuriously clustered point able to ruin down a straight segment estimate for the whole cluster.

Focusing on a cluster, the points belonging to the cluster are inliers. The points not belonging to the cluster are outliers. A first set of outliers are due to noise, or to non straight scene structures. A second group of outliers, the pseudo-outliers, corresponds with the other clusters. A robust method has to deal successfully with both type of outliers.

We present an adaptation for 2D range scans of the robust algorithm, ASSC. ASSC was introduced by Wang and Suter [3] for 3D images and fundamental matrix. It combines the advantages of two classical segmentation algorithms: RANSAC and LMedS. RANSAC is able to tolerate a high rate of outliers, but needs an estimate of the error scale for the clustering, a predefined scale is applied for all the clusters. LMedS can compute the scale of inliers from the data but only

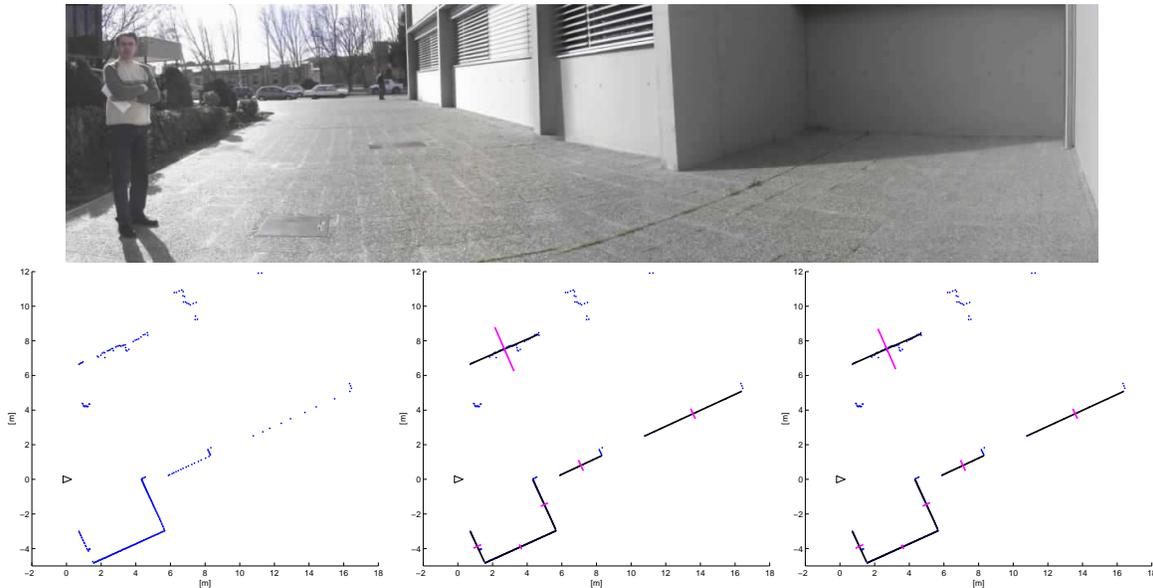


Fig. 1. This figures shows a real scan with different straight segments (pseudo-outliers) and unstructured data (outliers). Top shows a panoramic photo of the sensed area; notice the person, the ivy covered wall behind him on the left, and the concrete walls on the right. Bottom left shows the raw laser data. Bottom middle shows the manually detected segments. Bottom right shows the experimental results. The scale is plotted as red orthogonal dash in the middle of the segments, 30σ is used to magnify the scale in order to make it visible. Note the different scales computed for the ivy covered wall and for the concrete walls.

can deal with up to 50% of outliers. ASSC produces a data driven scale estimate per each cluster when segmenting, and is able to deal with high spurious rate as RANSAC.

ASSC is based on the mean shift search for maxima and minima in the probability density function (pdf) of the point error with respect to the straight segment defined by a cluster. The pdf is coded considering the scanner points as samples of the error distribution.

II. STATE OF THE ART

The problem of line extraction in 2D range scans has been widely studied in the robotics literature. Several algorithms have appeared with different properties and performance [1]. Here we introduced the most popular methods.

A. Line-building Algorithms

Line-building algorithms directly compute a suboptimal parametric solution of a multiple structure, exploiting the property that a single scan is an open chain of sorted points with no loops. Segmentation is done concatenating consecutive points that accomplish some heuristic line criteria. Usually, the criteria are very simple, like checking an error bound, being the fastest line detector algorithms. However, it requires a prior knowledge of inliers scale. Moreover, none probabilistic uncertainty model is assumed. Consequently, basic heuristic conditions work very well with simple pseudo-outliers, but crowded environments and gross outliers require complicated rules being highly dependable of the application.

Examples of these algorithms are Split and Merge [4] and Incremental [5], also called Line-Tracking. Essentially, *Split and Merge* recursively splits the original parameter model

(e.g.: a straight line between the first and last point), when the maximum residue is higher than a fixed threshold. Resulting lines are thereupon merged following the same rule, i.e: the maximum residue of two collinear segments is lower than the threshold. On the other hand, *Incremental Algorithm* starts with two close points (i.e: the first and second point), adding the next scan point to the end of the segment when the line criteria is validated. If the criteria is not achieved, then the current line is finished and a new line is started at the next point.

B. Hough Transform

The Hough transform [6] is based on a voting strategy to determine the *best fit* for a data subset. The main drawback of the method is that the parametric space must be discretized, consequently, the accuracy is highly affected in a real time application, since computational cost is $\mathcal{O}(nd)$ where n is the number of data points and d is the size of the voting grid.

Hough Transform resembles to our approach regarding there is no prior assumption about the error bounds or the level of spurious. In addition, as a single majority voting strategy, it can intrinsically deal with both outliers and pseudo-outliers.

C. Random Sampling Segmentation Algorithms

In this case, the aim of the algorithms is to find a suboptimal probabilistic model to classifying the data points and to separate inliers from outliers.

If we consider independent normally distributed inliers $x_i(\theta) \sim \mathcal{N}(0, \sigma)$, and assume that the scale estimate, or dispersion coefficient, corresponds to the estimated standard deviation σ . The inliers are taken to be those data points which

normalized squared residual follows a χ^2 distribution

$$\left(\frac{x_i(\theta)}{\sigma}\right)^2 \leq \chi_{1-\alpha,n}^2 \quad i = 1 \dots n \quad (1)$$

where the χ^2 distribution is a constant value for a fixed probability of false positives (i.e: $\alpha = 0.01$). Then, a scale factor σ is mandatory to classify the data.

The clustering step is made by random sampling, taken minimal subsets to compute several parametric models θ_i . The winning subset model $\hat{\theta}$ is that which maximizes an utility function U_θ . The optimal model is used to label the data using the scale parameter σ (see equation 1). These algorithms can intrinsically handle multiple structures using a recursive search through the remaining data. Finally, a TLS algorithm is applied to every cluster separately in order to correct the bias.

Considering that p is the size of a subset to compute a minimal parametric model (e.g: 2 for a segment). The number of samples to draw in order to find a subset with no outliers with probability P is [7]

$$m = \frac{\log(1 - P)}{\log[1 - (1 - \varepsilon)^p]} \quad (2)$$

where ε is the ratio of outliers.

In contrast to previously introduced algorithms the computational cost of random sampling approaches does not depend on the size of the data set but on the spurious rate.

1) *Least Median Squares (LMedS)*: The median estimator is probably the most extended robust estimator due to its simplicity and robustness when the ratio of inliers is higher than 0.5. In this case, the optimization criteria involve minimize the median of the squared errors. Using a voting analogy, it is similar to ask for an absolute majority.

Thus, the scale estimate is given by [7]

$$\hat{\sigma} = 1.4826 \left(1 + \frac{5}{n - p}\right) \sqrt{\text{med}_{\hat{\theta}} r^2} \quad (3)$$

where $\text{med}_{\hat{\theta}} r^2$ is the minimum median, n is the number of samples and p is the dimension of the parameter space.

The main problem of LMedS appear in scenes with multiple structures, because the level of pseudo-outliers is noticeably bigger than 0.5.

2) *Random Sampling Consensus: RANSAC*: RANSAC assumes the inliers are the largest cluster for a predefined scale, which can be an heuristic threshold or obtained from sensor calibration. Consequently, the optimization criteria consist of maximizing the number of inliers.

$$\hat{\theta} = \arg \max_{\hat{\theta}} n_{\hat{\theta}} \quad (4)$$

where $\hat{\theta}$ is the parameter estimate and $n_{\hat{\theta}}$ is the number of inliers. Provided a good scale estimate for every cluster, RANSAC is able to cope with large amounts of outliers. Nevertheless, the prior scale knowledge is sometimes unavailable or inaccurate. This *a priori* scale resembles to the Split and Merge and Incremental error bound.

Following in the voting analogy, this algorithm performs several referendums and it selects the option with biggest majority.

III. ASSC: A KERNEL-BASED SCALE ESTIMATOR

Adaptive Scale Sample Consensus (ASSC) [3] is a modification of RANSAC involving an adaptive scale estimation. The data driven scale estimate is computed using mean shift method [8]. In this case, the utility function takes into account both the number of inliers and the scale factor. Concluding the voting analogy, this algorithm merely requires a simple majority to achieve quorum.

Mean shift is an algorithm that can be used for clustering, in this case, collinear data. Given a random sample of the parametric model, we use the residual space to identify the clusters of data. The actual inlier data is the first cluster found with point error closer to 0. Subsequently, a LMedS algorithm is applied to the cluster in order to compute the final scale. In this case, LMedS converges to a solution because the cluster includes a single structure and few outliers.

A. Mean Shift Clustering

The mean shift paradigm [9] is a nonparametric method to find maxima and minima of an unknown pdf; the pdf is defined by a sample. In our case the random variable is the orthogonal error of a point with respect to a scene line segment, all the scanner points define the sample of the random variable.

$$M_h(x) \equiv H_h \frac{\hat{\nabla} f(x)}{\hat{f}(x)} \quad (5)$$

where $\hat{f}(x)$ is the density estimated and H_h is a bandwidth coefficient related to the *amount of information* [8]. It has been proved that mean shift vector points towards the direction of the maximum increase in the density. Consequently, the mode can be obtained using an iterative process

$$\mathbf{x}_{k+1} = \mathbf{x}_k + M_h(\mathbf{x}_k) \quad (6)$$

Mean shift vector computation is based on the kernel density estimation and Parzen windows theory. Given a set of samples $\mathcal{X} = \{X_1, \dots, X_n | X_i \in \mathbb{R}^d\}$, the kernel density estimator is defined as

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (7)$$

where $K(\cdot)$ is a *window* or kernel function of width h . We have selected the Epanechnikov kernel that yields the minimum mean integrated square error (MISE)

$$K^e(\xi) = \begin{cases} \frac{1}{2} c_d^{-1} (d+2) (1 - \xi' \xi) & \text{if } \xi' \xi \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where c_d is the volume of a d -dimensional hypersphere of unit radius (i.e: $c_1 = 2, c_2 = \pi$). If we define $\xi = (x - X_i)/h$ then, the Kernel function is an hypersphere $S_h(x)$ defined in the sampling space centered at x with radius h . For the Epanechnikov kernel, the amount of information H_h is

$$H_h^e = \frac{h^2}{d+2} \quad (9)$$

Consequently, using the Epanechnikov kernel, the mean shift vector can be rewritten as follows:

$$M_h^e(x) = \frac{1}{n_e} \sum_{\mathcal{X}_e} X_i - x \quad (10)$$

where $\mathcal{X}_e = \{X_{1..n_e} | X_i \in S_h(x)\}$ is the subset of n_e samples that fall within the hypersphere. As a result, the mean shift vector depends only on the samples and the scale, see (12). It can be proved that mean shift vector converges to the closest mode of the unknown pdf [10], which is the desired cluster for the current parametric solution. In addition, the same method can be used to find the closest valley to the mode, following the opposite mean shift vector

$$V_h^e(x) = -M_h^e(x) = x - \frac{1}{n_e} \sum_{\mathcal{X}_e} X_i \quad (11)$$

This valley (see algorithm 1) will be used to define the bounds for the data cluster. In our case, the selected interval is $[0, x_{valley}]$, because the residual data is always positive.

B. Adaptive Scale Sample Consensus (ASSC)

The main advantage of ASSC [3] is that the scale factor is computed for every cluster separately, instead of being a heuristic threshold.

Kernel density estimation [11] and, thereby, mean shift method [9], are based on the smoothing of the sample distribution. Wide kernels produce oversmoothed pdf, while narrow kernels yield peaked pdf. Consequently, the performance of the mean shift vector is related to the kernel width h . For instance, while searching the mode it is interesting to oversmooth the actual density function to avoid small local maxima during the gradient search. As a result, a oversmoothed bandwidth selector [10] has been chosen

$$h = \left[\frac{4}{3n} \right]^{\frac{1}{5}} S_q \quad (12)$$

where n is the total number of data and $S(q)$ is a coarse preestimation of the standard deviation, that is, the scale of inliers. In ASSC algorithm, Wang and Suter suggested using a generalization of the median estimator based in percentiles smaller than 50% [12]

$$S(q) = \frac{d_q}{\Phi^{-1} \left[\frac{1+q}{2} \right]} \quad (13)$$

where $q \in [0, 1]$ is the expected maximum ratio of inliers, d_q is the half-width of the shortest window including the fraction q of total residuals and $\Phi^{-1}[\cdot]$ is the argument of the normal cumulative density function. Note that $S(0.5)$ is the median estimator. We have tested during the experiments that the q value is not significant for the performance of line extraction. For instance, a good policy is to use a pessimistic assumption in the ratio of inliers. Wang and Suter [3] suggest $q = 0.2$, which is analogous to rely on the 20% of data points with minimum error, for the initial scale estimation used to define h in equation (12).

Thereupon, mean shift is applied to find the bounds of the cluster. Finally, the actual scale of every cluster is recomputed using LMedS algorithm, taking in to account only the data inside the cluster. Hence, we avoid the problem of LMedS with multiple structures.

The utility function should be directly proportional to the number of points n_k and inversely proportional to the dispersion of points σ_k . Therefore, the utility function is defined as:

$$U(\theta_k) = n_k / \sigma_k \quad (14)$$

In consequence, RANSAC is a particular case of ASSC for a fixed scale model. ASSC has the capability to handle multiple scale models even in a single scan.

Algorithm 1 ASSC algorithm

```

for iteration in 1 to m do
   $\theta_k = g(Y_p) | Y_p \sim \mathcal{Y}$  subset sample
   $\mathcal{X} = res(\mathcal{Y}, \theta_k)$ 
  compute  $h$ 
  repeat
    compute  $M_h(x_i)$ 
     $x_{i+1} = x_i + M_h(x_i)$ 
  until  $M_h(x_i) = 0 \Rightarrow x_i = x_{peak}$ 
  repeat
    compute  $V_h(x_i)$ 
     $x_{i+1} = x_i + V_h(x_i)$ 
  until  $V_h(x_i) = 0 \Rightarrow x_i = x_{valley}$ 
   $X_k = \mathcal{X} \in [0, valley]$ 
   $\sigma_k = 1.4826 \left( 1 + \frac{5}{n-p} \right) \sqrt{med X_k^2}$ 
   $U(\theta_k) = n_k / \sigma_k$ 
end for
 $\hat{\theta} = \arg \max U(\theta_k)$ 

```

Since ASSC has no restriction about scale, it needs a stop criteria to avoid selecting outlier data after having clustered all inliers. A basic test has been implemented to detect the directionality of data based on the ratio between segment length and estimated scale $(l_i / \sigma_i) \geq \beta$. In the subsequent experiments, $\beta = 10$.

IV. EXPERIMENTS

For the experimental results a data set has been collected using a mobile platform equipped with a SICK laser range finder along Ada Byron building, at the University of Zaragoza. The robot has been driven along 600 meters indoor and outdoor manmade environment. The environment is interesting due to the presence of different walls (concrete, tiles, glass, steel). For example, an important piece of the trajectory is done next to a ivy-covered wall (see pictures in figures 1 and 3) which increase the noise level, and hence the scale, of those features. There are also some blinds and grassy slopes which provide poor and noisy reflections. The architecture of the building presents challenging staircase-shaped walls with acute and obtuse angles (figure 3). Furthermore, there are some curved

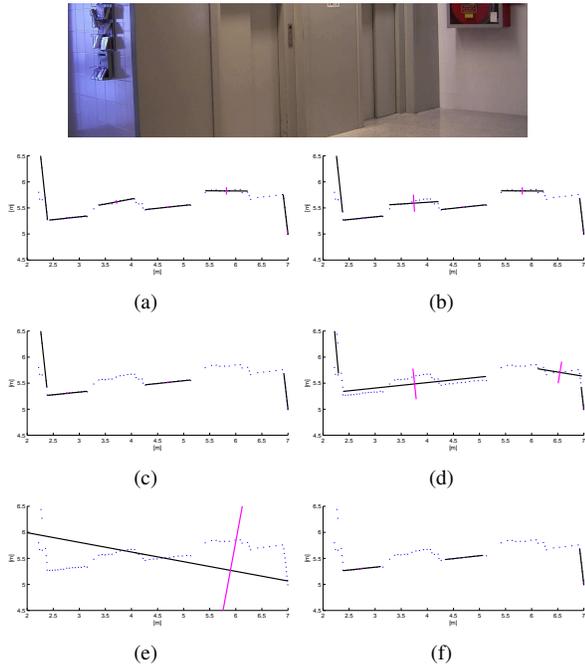


Fig. 2. Door segmentation a) Manual segmentation b) ASSC c) RANSAC with low noise model d) RANSAC with high noise model e) LMedS f) Split and Merge + LMedS

TABLE I

COMPARISON OF DETECTION PERFORMANCE IN SCALE ALGORITHMS.

Algorithm	False neg.	False pos.	Split segs	Merged segs
LMedS	32%	33%	38%	30%
S&M + LMedS	31%	3%	85%	0%
RANSAC high	16%	24%	19%	27%
RANSAC low	12%	11%	48%	5%
ASSC	10%	17%	19%	11%

elements like people, baskets and decorative elements, which should be detected as outliers during segmentation.

Some random scans and some specially difficult scans have been selected to perform a manual segmentation of data points. The whole experiment consist of 38 labeled scans and 190 features. A total least squares (TLS) estimated has been computed with the labeled data as a benchmark for the algorithms.

We contrast the reliability of ACCS algorithm versus RANSAC with two different predefined scales (high and low), LMedS and a combination of Split and Merge and LMedS, as proposed in [13]. The only constrains imposed to segmentation algorithm are the minimum number of inliers $n_{min} = 10$ and the maximum distance between points to break the segment $d_{max} = 1m$. The manual benchmark takes into account these constrains. See table I for a comparative in estimate performance. If the actual segment length differs more than 10% compared with the extracted segment, we classify the segment as split or merged respectively.

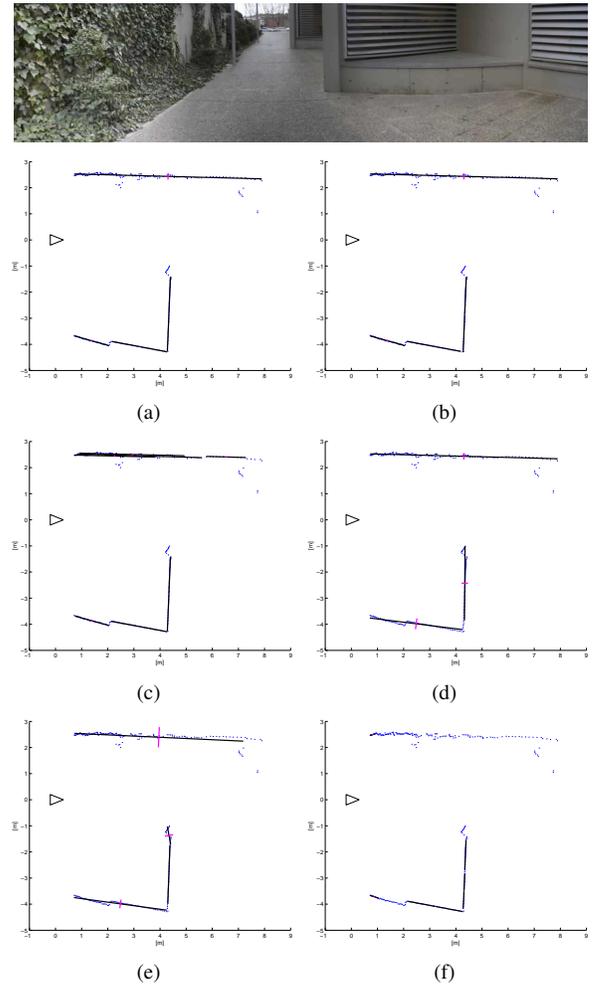


Fig. 3. Corridor segmentation. Up wall: ivy Down walls: concrete and blinds. For ASSC, ivy wall has bigger uncertainty that the concrete wall. a) Manual segmentation b) ASSC c) RANSAC with low noise model d) RANSAC with high noise model e) LMedS f) Split and Merge + LMedS

ASSC offers a good compromise between all false positive, false negative, wrong split and wrong merged segments. Next in performance is RANSAC, having the problem of finding a unique predefined scale for all the segments. RANSAC with high level noise fails in in the corners, where it is difficult to distinguish the end current line and the beginning of another line; usually, one or two pseudo-outliers are included in the RANSAC clusters. On the other hand, a low scale RANSAC splits noisy segments like ivy, providing an underestimate of the final segment. Finally, due to its low breaking point, LMedS algorithm is unstable in structures formed by several concatenated elements (i.e: more than 2) and Split and Merge discards or splits segments with outliers.

We have selected a specially difficult part of the experiment, which is very frequent in man made environments, to show the improvement of our approach: doors in a wall (see figure 2). This situation is very deceptive since features are almost collinear. In addition, some intermediate points appear in the doorframe, hindering segmentation.

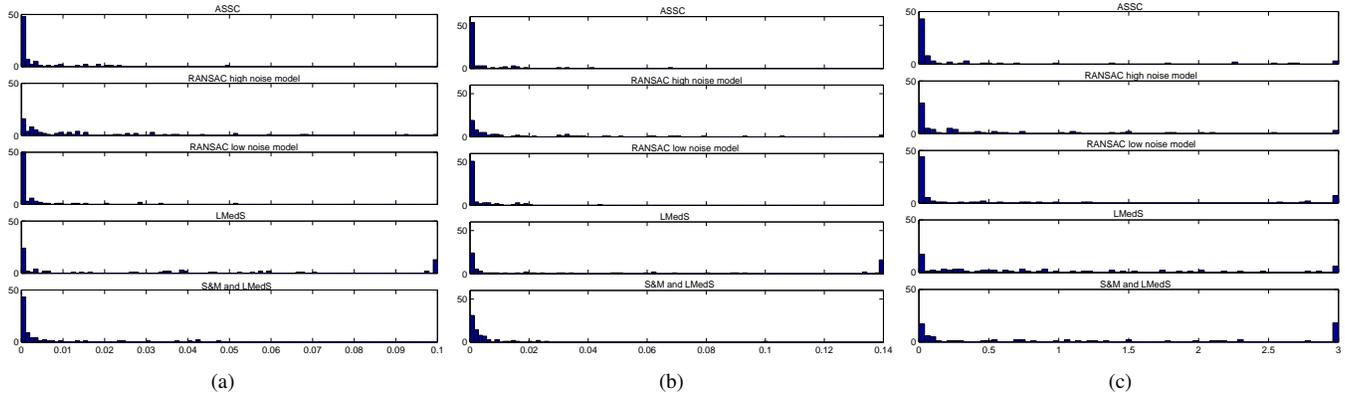


Fig. 4. Error distributions: a) lateral error, b) orientation error, c) length error

Usually, line extraction algorithms tend to include the doors as a part of the wall segment because the elements are collinear according to the error model or the algorithm can not deal with multiple structures. A simple approach would consist in adjust the error model. Nevertheless, this could lead to overconfidence and inaccuracies in other parts of the environment with more texture. In contrast, ASSC is able to distinguish between texture and multiple close structures, as shown in figure 2. Basically, our approach determine the best scale for each structure, which sometimes, is imposible to be determined previously.

Other interesting situation appears when moving through an outdoor corridor where on one side is a simple flat surface but covered with ivy and on the other side is a concrete wall, although with a complex structure (see figure 3).

Figure 4 shows the histograms of lateral, orientation and length error distribution of the features detected in the whole experiment. It is worth to notice that the histograms has been truncated to clearly represent the information. The last bin includes all error thereafter. Our approach also improve the accuracy of classical approaches since it is able to select a tight uncertainty for every segment.

V. CONCLUSION

This paper presents an algorithm for robust data segmentation involving multiple structures and high percentage of outliers. The main advantage is the capability to compute a data driven scale. This behavior improves, not only the segmentation but also the final accuracy. Furthermore, our approach does not require any previous calibration of the sensor or tuning of the parameters of the algorithm. This feature allows to reuse the code in different platforms, sensors and application. Furthermore, the algorithm has proved its robustness dealing with different objects in the same scene, for instance bumpy ivy-covered walls and flat concrete walls. Originally designed for 3D laser scanner, we have adapted and tested the ASSC algorithm for 2D range scan data. The method has been experimentally validated indoor and outdoor manmade environments for robotics applications, surpassing

current algorithms performance both in estimation and detection.

ACKNOWLEDGMENT

This research has been partially founded by the Spanish CICYT Spain under project DPI2003-07986.

REFERENCES

- [1] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, "A Comparison of Line Extraction Algorithms using 2D Laser Rangefinder for Indoor Mobile Robotics," in *Proc. 2005 IEEE Int. Conf. on Intelligent Robots and Systems*, 2005.
- [2] K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space," *Philosophical Magazine*, vol. 2, pp. 559–572, 1901.
- [3] H. Wang and D. Suter, "Robust adaptive-scale parametric model estimation for computer vision," *IEEE Trans. Pattern Analysis Machine Intell.*, vol. 26, no. 11, pp. 1459–1474, November 2004.
- [4] S. Horowitz and T. Pavlidis, "Picture segmentation by a tree traversal algorithm," *JACM*, vol. 23, no. 2, pp. 368–388, April 1976.
- [5] R. Taylor and P. Probert, "Range Finding and Feature Extraction by Segmentation of Images for Mobile Robot Navigation," in *Proc. 1996 IEEE Int. Conf. on Robotics and Automation*, 1996.
- [6] P. Hough, "Methods and means for recognising complex patterns," US Patent 3 069 654, 1962.
- [7] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*. John Wiley & Sons, 1987.
- [8] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Analysis Machine Intell.*, vol. 24, no. 5, pp. 603–619, 2002.
- [9] Y. Cheng, "Mean Shift, Mode Seeking, and Clustering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 8, pp. 790–799, August 1995.
- [10] M. Wand and M. Jones, *Kernel Smoothing*. Chapman & Hall, 1995.
- [11] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley Interscience, 2001.
- [12] K. M. Lee, P. Meer, and R. H. Park, "Robust Adaptive Segmentation of Range Images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, no. 3, pp. 200–205, February 1998.
- [13] P. Newman, J. Leonard, J. Neira, and J. Tardós, "Explore and Return: Experimental Validation of Real Time Concurrent Mapping and Localization," in *Proc. 1996 IEEE Int. Conf. on Robotics and Automation*, Washington D.C., May 2002, pp. 1802–1809.