



CC-RANSAC: Fitting planes in the presence of multiple surfaces in range data

Orazio Gallo^{a,*}, Roberto Manduchi^a, Abbas Rafi^b

^a University of California, Santa Cruz, United States

^b Canesta, Inc., United States

ARTICLE INFO

Article history:

Received 23 July 2009

Available online 3 November 2010

Communicated by B. Kamgar-Parsi

Keywords:

Robust fitting

RANSAC

Range data processing

Time-of-flight applications

ABSTRACT

Range sensors, in particular time-of-flight and stereo cameras, are being increasingly used for applications such as robotics, automotive, human-machine interface and virtual reality. The ability to recover the geometrical structure of visible surfaces is critical for scene understanding. Typical structured indoor or urban scenes are often represented via compositional models comprising multiple planar surface patches. The RANSAC robust regression algorithm is the most popular technique to date for extracting individual planar patches from noisy data sets containing multiple surfaces. Unfortunately, RANSAC fails to produce reliable results in situations with two nearby patches of limited extent, where a single plane crossing through the two patches may contain more inliers than the “correct” models. This is the case of steps, curbs, or ramps, which represent the focus of our research for the impact they can have on cars’ safe parking systems or robot navigation. In an effort to improve the quality of regression in these cases, we propose a modification of the RANSAC algorithm, dubbed CC-RANSAC, that only considers the largest connected components of inliers to evaluate the fitness of a candidate plane. We provide experimental evidence that CC-RANSAC may recover the planar patches composing a typical step or ramp with substantially higher accuracy than the traditional RANSAC algorithm.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Range sensors, in particular time-of-flight (TOF) and stereo cameras, are being increasingly used for applications such as robotics, automotive, human-machine interface, and virtual reality. The ability to recover the geometrical structure of visible surfaces (for example, using parametric models such as planar patches or other geometric primitives) is critical for scene understanding. For example, consider a sensory system for assisted back-up and parking (Glazduri, 2005; Paine et al., 2005; Backup systems, 2004; Hsu et al., 2007). To be really effective, such systems should be able to reason about the scene structure, identifying, for example, planar patches and discontinuities. In particular, they should robustly identify and localize structures such as curbs and ramps, like those of Fig. 1, as these are important features for safe parking.

A classical method for range analysis with this type of structures is to extract the dominant planar structures (for example, the “ground plane”), and to model the geometric feature as a composition of planar patches. Unfortunately, the presence of multiple planar structures at close vicinity and orientation, may impair detection of the dominant plane using classical methods (e.g., RANSAC (Fischler and Bolles, 1981)). Consequently, hazard detection

approaches which rely on detection of a dominant plane, such as the one of Hsu et al. (2007), may be adversely affected by the presence of curbs and ramps. An example is shown in Fig. 6 (c): rather than selecting one of the three possible planar patches forming the curb, RANSAC chose a plane intersecting all three. This type of error, which is by no means unusual (Stewart, 1997), may impair height measurements of the objects in the scene, since height is usually measured with reference to the ground plane.

This paper presents an improved algorithm for plane fitting, dubbed CC-RANSAC, shown to be more reliable than RANSAC in these situations. Whereas RANSAC uses the whole set of inliers to evaluate the fitness of a candidate plane, CC-RANSAC only considers the largest connected components of inliers at each iteration. This seemingly minor modification is in fact key to a substantial improvement in estimation accuracy, as evaluated with experiments in synthetic and real data from a TOF camera.¹

This contribution is organized as follows. We first review the major algorithms for range analysis as well as curb and step detection in Section 2. In Section 3 we describe our approach to plane fitting and evaluate it on synthetic data (Section 3.1). We then perform a thorough case study in Section 3.2. Finally, in Section 3.3 we present more results on real data.

* Corresponding author. Fax: +1 831 459 4046.

E-mail address: orazio@soe.ucsc.edu (O. Gallo).

¹ A shorter version of this paper was presented at the time-of-flight workshop held in conjunction with the conference of Computer Vision and Pattern Recognition.

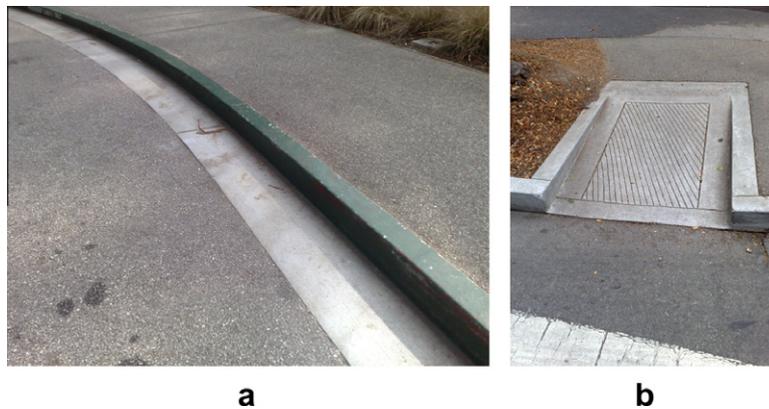


Fig. 1. An example of a curb (a) and of a ramp (b). These types of features must be identified for safe parking.

2. Background and previous work

2.1. Algorithms for range processing

Research on range analysis represents a vast body of work, encompassing Computer Vision, Robotics, and Computer Graphics. In the following we attempt a simple organization, with the purpose of providing some context and background for the proposed research. A simple categorization of range analysis algorithms may be drawn based on whether local or global descriptors are employed. Local descriptors include local surface normals (Lalonde et al., 2005; Mitra et al., 2004), ridges (Eberly et al., 1994), and discontinuities (Tang et al., 2004; Tong et al., 2001; Adams and Kerstens, 1998; Adams, 2001). Contiguous point sets with similar local descriptors may be clustered in space in order to identify extended regions. For example, chains of points with high curvature may form a curb line, and groups of adjacent points with the same normal may identify a planar patch. Only local analysis of the range data is required for this type of descriptors, which can therefore be computed very quickly. For the same reason, however, local descriptors are susceptible to measurement noise and missing measurements.

“Global” descriptors, on the converse, are parametric representations (typically planar or quadratic) of relatively large surface patches. All measurements in a patch contribute to the estimation of the parameters of the global descriptor. For example, if a set of measurements are known to be part of a plane, then simple linear regression (perhaps using principal component analysis, PCA) can provide the corresponding planar equation.

When the measurements are affected by “outliers” (data points that differ substantially from the standard noise model), robust procedures should be employed (Stewart, 1999; Meer, 2004). For example, M-estimators find the model parameters that minimize a cumulative “robust” loss function. With respect to the quadratic loss function used for standard linear regression, robust loss functions penalize more those samples that deviate heavily from the model. Possibly, the best known robust parametric estimators in Computer Vision are RANSAC (Fischler and Bolles, 1981) and the Hough transform (Illingworth and Kittler, 1988). Both can be seen as particular instances of M-estimators (Stewart, 1999). Another popular robust estimation method is the Least Median of Squares (LMedS) (Rousseeuw and Leroy, 1987) and its variants, which include the Least Kth of Squares (LKS). It can be shown that LMedS and LKS are instances of so-called S-estimators, which are a particular case of M-estimators (Chen et al., 2001). Another approach to dealing with outliers is to explicitly model them as uniformly distributed. This assumption is at the basis of the MLESAC algorithm (Torr and Zisserman, 2000).

An important parameter of robust estimators is the “scale”, ε , at which they operate. Intuitively, those points that are at a distance larger than ε from the estimated plane are considered “outliers”; the remaining points are “inliers”. Clearly, the scale depends on the variance of the inliers, usually modeled as normally distributed. The choice of scale may critically affect the performance of an estimator. A number of solutions to the scale estimation problem exists, including joint estimation with the model parameters (Huber, 1981), minimum unbiased scale estimation (MUSE (Miller and Stewart, 1996)), adaptive least Kth order square estimation (ALKS (Lee et al., 1998)) and modified selective statistical estimation (MSSE (Bab-Hadiashar and Suter, 1998)). When the variance of the inliers is not constant (heteroscedastic data), then more complex robust algorithms should be used (Subbarao and Meer, 2005).

In general, a given planar patch occupies only a finite portion in the image, with other, competing planar regions present as well. There are three main approaches for the simultaneous segmentation and estimation of planar regions in the same image. The first approach, which we use for the experiments in this paper, is to simply use a robust estimator to extract a “dominant” planar region, by considering all the remaining points (including any other planar regions) as outliers. After finding the planar region and removing the inliers, the operation is repeated on the remaining points, until no more sizable planar structures can be found. This algorithm is simple and intuitive, however, the presence of multiple structures may impair the estimation of individual planar patches, especially if the scale is not estimated correctly. This phenomenon was studied in detail (Stewart, 1997; Chen et al., 2001).

The second approach to multiple model estimation is to run a simultaneous, concurrent optimization over all planes visible in the image. This can be obtained using the Expectation-Maximization algorithm (an iterative technique akin to K-means clustering) (Liu et al., 2001; Triebel et al., 2005) or the recently developed Generalized PCA algorithm (Vidal et al., 2005). In this case, each plane is represented explicitly, rather than resorting to the notion of “outlier” with respect to a dominant structure. While intuitively more appealing, this approach requires the joint estimation of the (unknown) number of planar surface elements in the scene, an operation that often proves challenging (McLachlan and Peel, 2000).

The third family of algorithms is based on region growing (Besl and Jain, 1988; Taubin et al., 1991). Starting from some “seed” points or regions, homogeneous patches are grown concurrently by adding neighboring points consistent with the model. Regions that have similar models can then be merged together. Both region growing and merging can be performed using robust criteria (Boyer et al., 1994; Koster and Spann, 2000). Region growing is a simple

and fast algorithm, but relies on the selection of good seed points, which may be difficult to obtain, especially when planar patches of interest occupy only a small portion of the image. Note that region growing can also be used as an initial step for subsequent robust parameter estimation (Unnikrishnan and Hebert, 2003).

2.2. Curb and step detection using range data

Curb detection over short distances for safe driving has been demonstrated at CMU with a laser striper (Aufreire et al., 2003). The problem with a fixed laser striper is that the viewing geometry is very limited, while our task requires the ability to detect features over a rather wide field of view.

Se and Brady (1997) used a stereo camera pair to detect curbs and steps. They detect candidate curbs by finding clusters of lines in an image using the Hough Transform. Then, in order to classify a curb line as a step-up or step-down, they compute the ground plane parameters of the two regions separated by the curb line. This allows one to precisely estimate the height of the curb.

The work of Turchetto and Manduchi (2003) combined stereo and visual information to find step edges. The idea behind this approach is that a curb's edge usually generates a brightness edge in the image, and thus, in the neighborhood of the projected curb's edge, the elevation gradient and the brightness gradient are expected to both have high values and to be aligned. Accordingly, detection is based on a weighted Hough Transform on brightness edge points, with weights values proportional to the scalar product of the brightness gradient and of the depth gradient in the image.

This idea is pushed further in the work of Lu and Manduchi (2005). For each surface element within a certain distance from an estimated ground plane, a surface curvature measure is computed on the range data, characterizing the likelihood that the point belongs to a curb or step edge. Segments in the image that are characterized by high brightness gradient (edges) and high surface curvature are extracted by means of a weighted Hough transform. Finally, these segments are reprojected back into the 3-D scene. The algorithm produced the endpoint of a 3-D segment representing the curb edge, and could be used also to characterize staircases.

More recently, Pradeep et al. (2008) proposed another stereo-based system for curb detection that is based on plane fitting. Tensor voting is used to calculate consistent normals at each data point, which allow for clustering into planar patches.

3. Regression and CC-RANSAC

As discussed previously, in the presence of curbs or small steps, dominant plane detection may produce unsatisfactory results. This was noted, for example, by Lu and Manduchi (2005), where it was shown that the estimated “ground plane” was not reliable enough for detection of small steps. In order to understand this behavior, it may be useful to quickly review some basic concept of planar estimation.

Planar regression from a set of 3D point seeks for a plane \mathcal{P} that minimizes some measure of observed “fitness” to the data points. If d_i is the Euclidian distance of the i th data point to a candidate plane \mathcal{P} , different measures of the fitness $o(\mathcal{P})$ can be considered:

$$\begin{aligned} o &= -\sum_i d_i^2 \quad (\text{LS}) \\ o &= -\text{median}\{d_i^2\} \quad (\text{LMedS}) \\ o &= |I_c(\mathcal{P})| \quad (\text{RANSAC}) \end{aligned} \quad (1)$$

where $I_c(\mathcal{P})$ is the set of *inliers* (i.e., data points with $d_i \leq \varepsilon$ for a given threshold ε) and $|I|$ represents the cardinality of the set I . In the

Least Squares approach (LS), the plane \mathcal{P} with maximum fitness o can be found in closed form. In the other two cases, random sampling can be used for minimization. In general, when the variance of the noise is known, at least approximately, RANSAC is preferable to LMedS due to its lower computational cost. Both LMedS and RANSAC are superior to LS when outliers are expected or, as in the scenarios considered here, when multiple planar models are present in the scene.

However, as mentioned earlier, even RANSAC (or LMedS) may provide poor results when the scene contains two or more planar patches at short distance from each other. This phenomenon was studied at length by Stewart (1997). This is not a defect of sampling: rather, the proposed measure of fitness o is not adequate, in the sense that the planes representing different surfaces in the scene do not necessarily produce large values of fitness o . This is shown by way of example in Fig. 6. In this case, the plane that maximizes RANSAC fitness o (i.e., the plane receiving the highest number of supporting inliers) is shown in a Fig. 6 (c). This plane straddles across the two planes representing the top and bottom surfaces of a curb. Similar results are obtained using the LMedS criterion.

Thus, even the robust fitness measures in (1) fail to correctly identify the individual visible planar components. We argue that the main problem with such measures is that they neglect the spatial coherence typically exhibited by inlier points. Accordingly, we propose a modification of the RANSAC algorithm, dubbed CC-RANSAC, by defining the following measure of fitness:

$$o = |I_c(\mathcal{P})| \quad (\text{CC-RANSAC}) \quad (2)$$

where $I_c(\mathcal{P})$ is the largest connected component of inliers with 8-neighbor topology inherited from the image grid. This idea embeds the observation that data points that are the inliers of a “correct” plane cluster contiguously in space, whereas a plane straddling across two planar patches typically produces two disconnected sets of inliers (see Fig. 6 (c)).

Using I_c for evaluating the fitness of a candidate plane ensures that only the inliers from a single planar patch will contribute to this measure. This is indeed the case for Fig. 6 (d), where the red points represents the inliers belonging to I_c for the same plane as in Fig. 6 (c).

3.1. Comparative performance assessment – synthetic data

In order to compare the performance of RANSAC and CC-RANSAC quantitatively, we first consider a synthetic data set with noisy 3-D points generated from a model of a step. This allows us to test the algorithm under a wide variety of controlled conditions. A range imaging system is assumed to collect data from two planar patches (each providing 150 by 50 measurements on a regular grid with point spacing of 1 unit along each axis). The two patches, which are separated by a distance of h units, are seen from above under orthography. The measurements are corrupted by Gaussian noise with standard deviation of σ units.

Our initial experiments computed robust planar regression with different values of the distance between the planar patches, h . Only one plane is estimated at each time. Since the two patches have the same number of measurements, ideal robust regression would produce a plane modeling either patch. In order to measure the discrepancy between the plane \mathcal{P} computed by the algorithm and the fitted patch, we compute the average square distances $\{\bar{d}_1^2, \bar{d}_2^2\}$ of the points of the two patches to the plane \mathcal{P} . Then, we define the regression error as $e(\mathcal{P}) = \min(\bar{d}_1, \bar{d}_2)$. In each experiment, we first fix the number N of random samples used for RANSAC or CC-RANSAC. More precisely, a set of N non-collinear triplets of measurements are sampled without replacement from the data pool; the fitness of the plane \mathcal{P} identified by each triplet is



Fig. 2. Synthetically generated data representing a curb with height h equal to 5 units, with added Gaussian noise (in the vertical direction) with standard deviation σ equal to 1 unit.

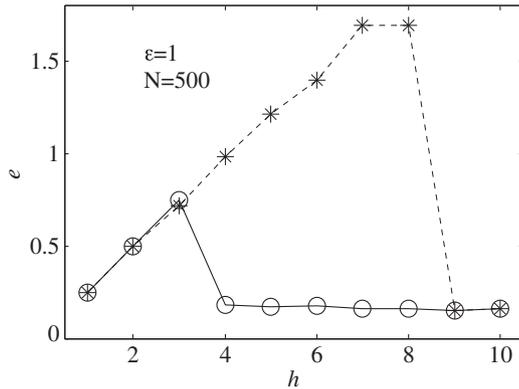


Fig. 3. Median regression error e as a function of the step height h using RANSAC (stars) and CC-RANSAC (circles).

computed using the RANSAC and the CC-RANSAC criteria; the best fitting plane is chosen for both cases, and a final least squares regression based on all the inliers is computed. This procedure is repeated for 500 times, each time with a new set of N triplets.

The median value e of regression error over the 500 experiments is used for the plots in the following figures (see Fig. 2).

Fig. 3 shows the median regression error e when the distance h between the two planar patches is varied. The value ϵ of the threshold for the inliers (as defined in Section 2.1) is set equal to 1 (i.e., equal to the standard deviation of noise). $N = 500$ random samples are used for both algorithms. The most noteworthy characteristic of the measurements in Fig. 3 is that, for h ranging between 4 and 8, RANSAC produces a relative large error, which drops to a low value for $h > H$, where H is a break-down value that in this case is approximately equal to 8. The reason for this behavior is that for small values of h , RANSAC produces planes that straddle between the two patches. When the patches are far enough from each other (relative to the measurement noise), RANSAC can produce stable and robust results, reliably fitting either patch. The benefit of CC-RANSAC is that the break-down value H is reduced from 8 to 3. In other words, CC-RANSAC allows for planar fitting in a wider range of step heights than RANSAC for the step considered in these experiments.

Next, we look at the performance of both algorithms when the inliers threshold ϵ is changed. As mentioned earlier, a measurement point is considered an inlier with respect to a candidate plane \mathcal{P} when its distance to \mathcal{P} is less than ϵ . Robustness to incorrect (or “mismatched”) values of ϵ is important, since the actual standard deviation of noise σ is not always known with precision. Fig. 4 (a) shows the median regression error e for a step of height $h = 5$ as ϵ is changed between 0.25 and 3 (while the standard deviation of noise σ remains equal to 1). It is seen that RANSAC yields basically the same (large) regression error, regardless of ϵ . CC-RANSAC produces reliable results for ϵ between 0.5 and 1.25. For larger values of ϵ , it matches the results of RANSAC. This is not surprising: for large enough values of ϵ , an incorrect plane straddling across the two patches will produce a large number of connected inliers.

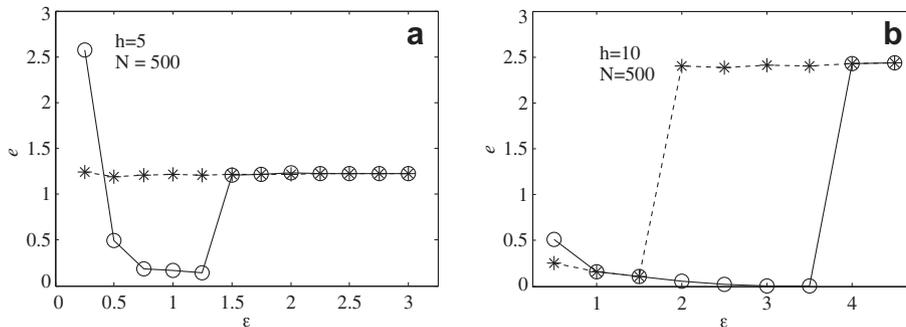


Fig. 4. Median regression error e as a function of inlier threshold ϵ using RANSAC (stars) and CC-RANSAC (circles) for two different values of step height ($h = 5$ and $h = 10$).

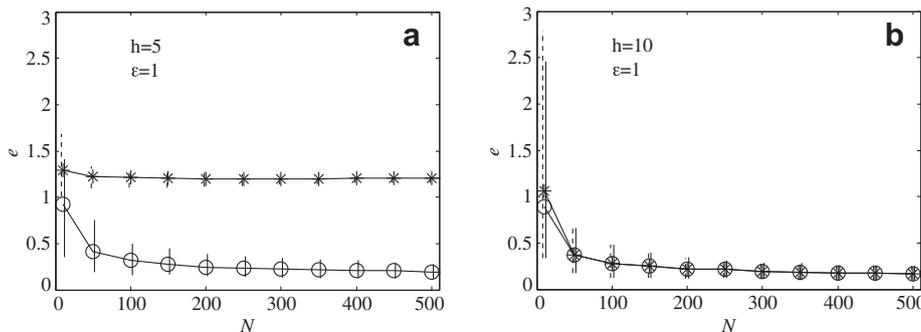


Fig. 5. Median regression error e as a function of the number of random samples N using RANSAC (stars) and CC-RANSAC (circles) for two different values of step height ($h = 5$ and $h = 10$). The bars show the 10- and 90-percentiles of the error distributions.

The only time CC-RANSAC performs worse than RANSAC is for very small values of ε . The reason in this case is that only very small connected component are formed, which cannot provide reliable support for the correct plane.

Fig. 4 (b) shows results from a similar experiment, but this time for a step with height h equal to 10 units. As seen in Fig. 3, at this step height RANSAC gives good results when ε is set to 1. When ε takes values larger than 1.5, though, RANSAC produces incorrect fitting planes with large median error. Remarkably, the breakdown point for CC-RANSAC is quite larger: only for ε larger than 3.5 does CC-RANSAC start behaving like RANSAC.

Finally, in Fig. 5 we show the median, along with the 10- and 90-percentile, of the regression error as a function of the number N of random samples used by the algorithms. When step height is set to 5 (with $\sigma = \varepsilon = 1$), CC-RANSAC produces relatively stable results for $N \geq 100$. Note that the median error for RANSAC remains basically constant, giving experimental evidence to the fact that incorrect regression is not a consequence of poor sampling in this case. For $h = 10$, the two algorithms perform substantially as well when N is changed.

3.2. Comparative performance assessment – TOF measurements

Here we consider the measurements shown in Fig. 6 (a), acquired by a Canesta TOF camera in front of a curb, as a study case. The goal is to find the prominent planar patch, shown in red in Fig. 6 (a). (Although there are two more planar patches visible in the scene, the one shown in red in Fig. 6 (b) has the largest number of data points.)

Even when the ground truth plane \mathcal{P}_0 is available, it is impossible (or at least unpractical) to label each data point as belonging to a particular planar patch. This means that the regression error measure proposed in the previous section cannot be used here. We thus define a different “goodness” measure, that does not require knowledge of which patch each point belongs to. Given a candidate plane \mathcal{P} , we define its quality $q(\mathcal{P})$ as the number of inliers of \mathcal{P} that are also inliers of the “ground truth” plane \mathcal{P}_0 , normalized by the number of inliers of \mathcal{P}_0 :

$$q(\mathcal{P}) = |I(\mathcal{P}) \cap I(\mathcal{P}_0)| / |I(\mathcal{P}_0)| \quad (3)$$

Note that $q(\mathcal{P}) = 0$ when \mathcal{P} is far enough from the planar patch, and $q(\mathcal{P}) = 1$ when \mathcal{P} coincides with \mathcal{P}_0 . Hence, q seems like an appropriate and simple to compute measure for describing how well a given plane fits the planar patch.²

Our first step is to compute the statistical correlation between fitness o of a candidate plane and its quality q . More precisely, we estimate the joint probability density function (pdf) of q and o , $f_{q,o}(q, o)$ by sampling the space of possible planes, where for each plane sample \mathcal{P} , o is set equal to either $|I_e(\mathcal{P})|$ or to $|I_c(\mathcal{P})|$ based on the data of Fig. 6 (a). Note that, although the space of candidate planes is discrete (since each plane is determined by a triplet of measured points), we make the simplifying assumption that it is continuous in our analysis. The two joint pdfs, computed using the Parzen window method from a set of 5000 random sample planes, are shown in Fig. 7; these plots reveal that, for both choices of fitness, the joint pdf of q and o is characterized by two main “ridges”, corresponding to two different clusters of planes.

We now show how $f_{q,o}(q, o)$ can be used to evaluate the expected performance of RANSAC or CC-RANSAC. More precisely, let q_N be the random variable describing the quality of the plane chosen by either algorithm after N iterations (where each iteration corresponds to a randomly selected candidate plane). If $\{o_n\}$ are the

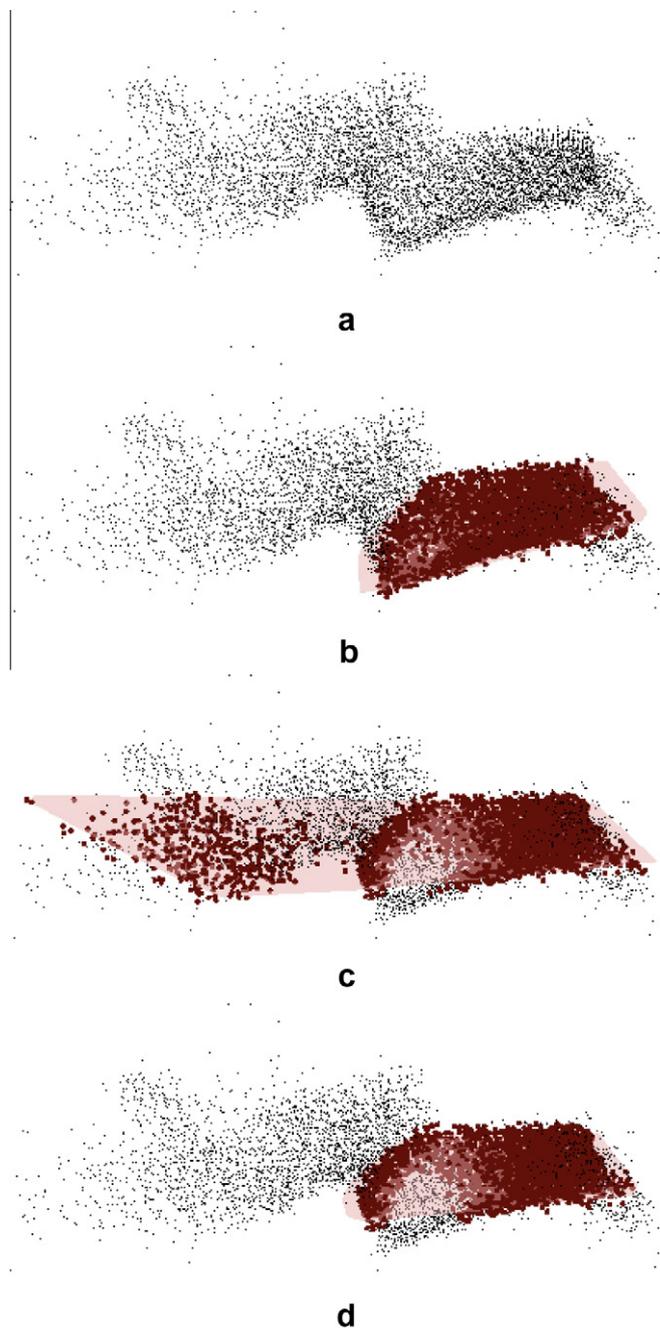


Fig. 6. (a): Range data collected in front of a curb. (b): Optimal planar fit to the lower planar patch (inliers with respect to the plane are shown in red). (c) Incorrect planar fit (\mathcal{P}) and inliers in $I_e(\mathcal{P})$. (d) Same plane as in (c) but with inliers in $I_c(\mathcal{P})$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

measured fitness values of the N candidate planes, then each algorithm chooses the plane \mathcal{P} with $o(\mathcal{P}) = \bar{o}$, where $\bar{o} = \max\{o_i\}$.

The pdf of q_N can be found as follows:

$$f_{q_N}(q) = \int_{-\infty}^{\infty} f_{q_N|o}(q|o) f_o(o) do \quad (4)$$

given that q_N represents the quality of the plane with the highest fitness measure, it is clear that $f_{q_N|\bar{o}}(q|\bar{o}) = f_{q|o}(q|\bar{o})$. The pdf of \bar{o} can be easily derived based on the fact that the samples are drawn independently:

$$f_{\bar{o}}(o) = N f_o(o) F_o^{N-1}(o) \quad (5)$$

² Note that, as opposed to the regression error e used in the previous section, a high value of the quality q indicates good algorithm performance.

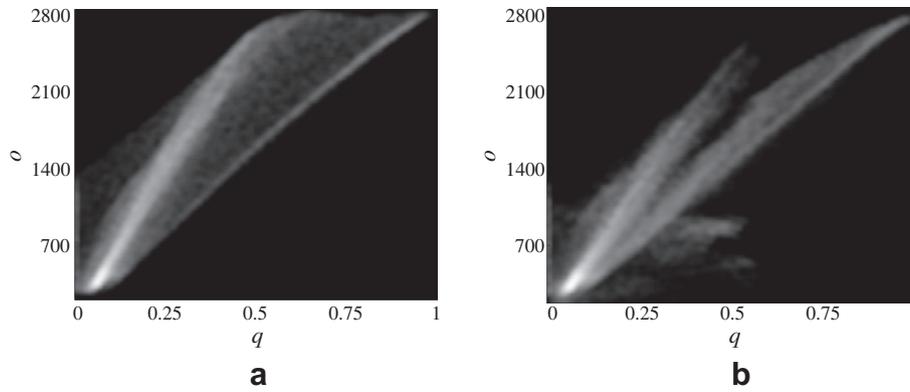


Fig. 7. Graphical representation of the joint densities $f_{q,o}(q,o)$ for the data of Fig. 6 (a) with (a) $o = |I_e(\mathcal{P})|$ and (b) $o = |I_{cc}(\mathcal{P})|$.

where $F_o(o)$ is the cumulative distribution function (cdf) of o :

$$F_o(o) = \int_{-\infty}^o f_o(u) du \quad (6)$$

All of those quantities are easily computed by numerical integration starting from $f_{q,o}(q,o)$.

Fig. 8 shows the pdf $f_{q_N}(q)$ for RANSAC and CC-RANSAC for different numbers N of iterations. It is interesting to note that RANSAC yields a bimodal distribution: since the two planes of Fig. 6 (a) and (c) both receive good inlier support, the algorithm may choose one or the other with almost the same likelihood (although the incorrect plane receives higher probability mass as the number of iterations increases). RANSAC-CC, instead, yields a unimodal distribution that is peaked around a high quality value, meaning that it almost invariably chooses a plane that is close to the optimal one. Even if only a limited amount of iterations (e.g., 100) is used, the chosen plane is likely to have a good quality value.

3.3. Planar fitting examples for curbs and ramps

This section presents a few experimental results using CC-RANSAC, in order to highlight the potential of this approach for curb and ramp detection. These examples are of interest for automotive applications, such as safe parking systems. Note that, although we only present results for ramps and steps, these patterns can be regarded as the building blocks for virtually any structure that only comprises planar surfaces. For example, an indoor staircase, which could be useful to detect for autonomous or semi-autonomous navigation (e.g., for the assisted control of a motor wheelchair (Murarka et al., 2006)), can be seen as a sequence of steps. The

crucial aspect, regardless of the type of structure, is that enough datapoints support each planar patch; in the case of autonomous navigation, for example, a staircase would probably be seen as a ramp from a distance and the steps would become visible as the robot, or the wheelchair, approaches it.

In Fig. 9, as well as in Fig. 6, inliers are represented with thick points, with color indicating to the plane they are closest to. For each fitting plane, we show the convex hull of its closest inliers, projected onto the planes.

Fig. 9 (a) shows the three best fitting planes to the data of Fig. 6. After the dominant plane has been found, the corresponding inliers are removed from the data, and the operation is repeated until a maximum number of planes is found, or the highest planar fitness for the remaining point is below a certain threshold.

Fig. 9 (b) shows the result to a similar curb taken from a larger distance. In this case, only two fitting planes were found. Note that the fit is pretty good, in spite of the planar patches being close to each other.

Examples of ramp modeling are shown in Fig. 9 (c) and (d). In particular, Fig. 9 (d) is based on measurements taken of the ramp shown in Fig. 1 (b). The red planar patch corresponds to the descending concrete surface in the ramp; the blue patch represents the asphalt surface at the bottom of the ramp; while the green patch corresponds to the surface covered in soil to the left of the ramp. Note that, contrary to what one would hope, the green and the blue patches do not intersect. This is due to the fact that the measurements supporting the green patch are biased by the presence of a tree stump, visible near the left edge of Fig. 1. Nonetheless, the algorithm is shown to produce very good planar fits to the different elements of the scene, which may enable further reasoning and recognition.

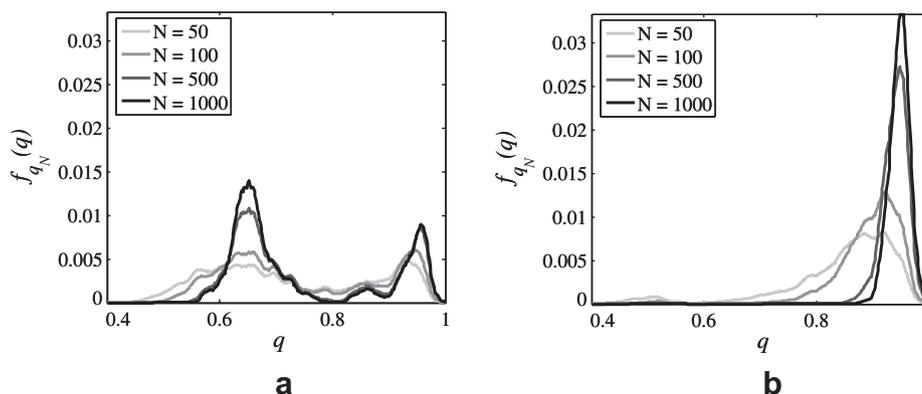


Fig. 8. Plots of the pdf $f_{q_N}(q)$ of the quality of the plane chosen in the case of Fig. 6 using (a) RANSAC and (b) CC-RANSAC with a variable number N of iterations.

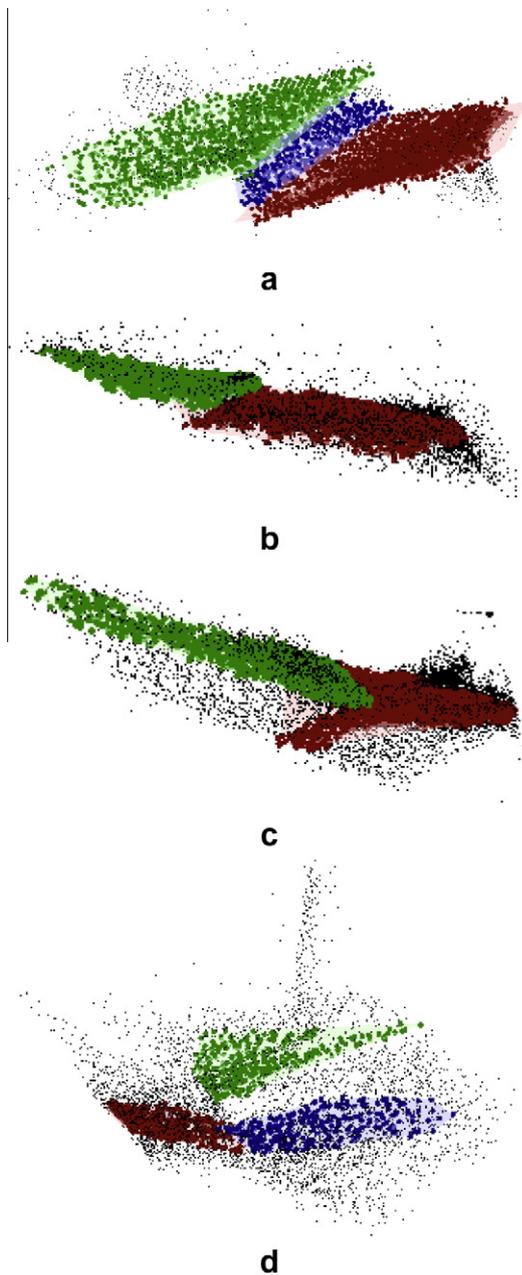


Fig. 9. Some experimental results with curbs ((a) and (b)) and ramps ((c) and (d)).

4. Conclusions

The ability to accurately measure the geometry of the scene enables systems relying on range sensors, for instance TOF sensors, to recognize features that are critical for many applications, such as safe parking. In this paper we concentrate on situations where this task can be impaired by the presence of multiple planar structures as in the case of curbs and ramps, a situation that is extremely common in urban environments. These are particularly challenging features that cannot be reliably detected using conventional ultrasound and microwave sensors.

In order to describe the geometry of a curb or of a ramp, we perform robust fitting to the different visible planar patches. We have shown that the popular RANSAC algorithm may fail in the case of a shallow curb; this result is in agreement with previous work by Stewart (1997). In order to deal with these situations, we propose a new algorithm, CC-RANSAC, that uses only the largest connected

component of inliers to evaluate the fitness of a candidate plane. This seemingly minor modification may in fact yield substantially better fits than RANSAC.

A critical analysis of CC-RANSAC brings a consideration to light. The assumption that inliers cluster together into one large connected component, although intuitively correct, needs to be investigated further. It is clear that the size of the largest connected component depends on the distribution of the distances d_i of the data points to the candidate plane as well as on the chosen threshold ϵ . If ϵ is too small, only isolated inlier clusters will form, as shown by our experiments of Section 3.1. If ϵ is too large, clusters of inliers corresponding to different planar patches may end up connecting with each other.

In future work we will investigate methods to increase the robustness of this approach by considering different ways to cluster the inliers. For example, one could use the isophotic metric (Pottmann et al., 2004), that combines Euclidean distance and distance between normals. This could help in situations with “holes” in the range data, which are liable to create multiple connected components where only one connected component is expected.³

References

- Adams, M.D., 2001. On-line gradient based surface discontinuity detection for outdoor scanning range sensors. In: Proc. IEEE/RSJ Internat. Conf. on Intell. Robots and Systems (IROS '01), vol. 3, pp. 1726–1731.
- Adams, M., Kerstens, A., 1998. Tracking naturally-occurring indoor features in 2-D and 3-D with lidar range amplitude data. Internat. J. Robotics Res. 17 (9), 907–923.
- Aufriere, R., Mertz, C., Thorpe, C., 2003. Multiple sensor fusion for detecting location of curbs, walls, and barriers. In: Proc. IEEE Intell. Vehicles Symp., pp. 126–131.
- Bab-Hadiashar, A., Suter, D., 1998. Robust range segmentation. In: Proc. 14th Internat. Conf. on Pattern Recognition, vol. 2, pp. 969–971.
- Backup systems, 2004. Consumer Reports 69 (10), 19–20.
- Besl, P.J., Jain, R.C., 1988. Segmentation through variable-order surface fitting. IEEE Trans. Pattern Anal. Machine Intell. 10 (2), 167–192.
- Boyer, K., Mirza, M., Ganguly, G., 1994. The robust sequential estimator: A general approach and its application to surface organization in range data. IEEE Trans. Pattern Anal. Machine Intell. 16 (10), 987–1001. <http://doi.ieeecomputersociety.org/10.1109/34.329010>.
- Chen, H., Meer, P., Tyler, D.E., 2001. Robust regression for data with multiple structures. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2001), vol. 1, pp. 1069–1075.
- Eberly, D., Gardner, R., Morse, B., Pizer, S., Scharlach, C., 1994. Ridges for image analysis. J. Math. Imaging and Vision 4 (4), 353–373. <http://dx.doi.org/10.1007/BF01262402>.
- Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Comm. ACM 24 (6), 381–395. <http://doi.acm.org/10.1145/358669.358692>.
- Glazduri, V., 2005. An investigation of the potential safety benefits of vehicle backup proximity sensors. In: Proc. Internat. Tech. Conf. on Enhanced Safety Vehicles.
- Hsu, S., Raffi, A., Hirvonen, D., 2007. Object detection and tracking using an optical time-of-flight range camera module for vehicle safety and driver assist applications. In: Proc. SAE World Cong. Exhibition.
- Huber, P., 1981. Robust Statistics. Wiley, New York.
- Illingworth, J., Kittler, J., 1988. A survey of the Hough transform. Comput. Vision Graphics Image Process. 44 (1), 87–116. [http://dx.doi.org/10.1016/S0734-189X\(88\)80033-1](http://dx.doi.org/10.1016/S0734-189X(88)80033-1).
- Koster, K., Spann, M., 2000. MIR: An approach to robust clustering-application to range image segmentation. IEEE Trans. Pattern Anal. Machine Intell. 22 (5), 430–444.
- Lalonde, J.F., Unnikrishnan, R., Vandapel, N., Hebert, M., 2005. Scale selection for classification of point-sampled 3D surfaces. In: Fifth Internat. Conf. on 3-D Digital Imaging and Modeling (3DIM 2005), pp. 285–292.
- Lee, K.-M., Meer, P., Park, R.-H., 1998. Robust adaptive segmentation of range images. IEEE Trans. Pattern Anal. Machine Intell. 20 (2), 200–205.
- Liu, Y., Emery, R., Chakrabarti, D., Burgard, W., Thrun, S., 2001. Using EM to learn 3D models with mobile robots. In: Proc. Internat. Conf. on Machine Learning (ICML).
- Lu, X., Manduchi, R., 2005. Detection and localization of curbs and stairways using stereo vision. In: Proc. IEEE Internat. Conf. on Robotics and Automation (ICRA 2005), pp. 4648–4654.
- McLachlan, G., Peel, D., 2000. Finite Mixture Models. Wiley-Interscience.
- Meer, P., 2004. Robust techniques for computer vision. In: Medioni, G., Kang, S.B. (Eds.), Emerging Topics in Computer Vision, first ed. IMSC Press Multimedia Series, Prentice Hall, Chapter 4.

³ We thank the anonymous reviewer for this suggestion.

- Miller, J.V., Stewart, C.V., 1996. MUSE: Robust surface fitting using unbiased scale estimates. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '96), pp. 300–306.
- Mitra, N.J., Nguyen, A., Guibas, L., 2004. Estimating surface normals in noisy point cloud data. *Internat. J. Comput. Geometry Appl.* 14 (4–5), 261–276.
- Murarka, A., Modayil, J., Kuipers, B., 2006. Building local safety maps for a wheelchair robot using vision and lasers. In: Proc. Third Canadian Conf. on Comput. and Robot Vision (CRV'06), Washington, DC, USA, p. 25. <http://dx.doi.org/10.1109/CRV.2006.20>.
- Paine, M., Henderson, M., 2005. Devices to assist in reducing the risk to young pedestrians from reversing motor vehicles, technical Specification 149. Roads and Traffic Authority, New South Wales, Australia (October).
- Pottmann, H., Steiner, T., Hofer, M., Haider, C., Hanbury, A., 2004. The isophotic metric and its application to feature sensitive morphology on surfaces. In: Proc. ECCV 2004, pp. 560–572.
- Pradeep, V., Medioni, G., Weiland, J., 2008. Piecewise planar modeling for step detection using stereo vision. In: *Comput. Vision Appl. for the Visually Impaired (CVAI'08)*, Marseille, France.
- Rousseeuw, P., Leroy, A., 1987. *Robust Regression and Outlier Detection*. John Wiley & Sons, Inc., New York, NY, USA.
- Se, S., Brady, J., 1997. Vision-based detection of kerbs and steps. In: Proc. British Machine Vision Conf. (BMVC '97), pp. 410–419.
- Stewart, C.V., 1997. Bias in robust estimation caused by discontinuities and multiple structures. *IEEE Trans. Pattern Anal. Machine Intell.* 19 (8), 818–833.
- Stewart, C.V., 1999. Robust parameter estimation in computer vision. *SIAM Rev.* 41 (3), 513–537.
- Subbarao, R., Meer, P., 2005. Heteroscedastic projection based m-estimators. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '05), vol. 3, pp. 38–38.
- Tang, F., Adams, M., Ibanez-Guzman, J., Wijesoma, W.S., 2004. Pose invariant, robust feature extraction from data with a modified scale space approach. Proc. IEEE Internat. Conf. on Robotics and Automation (ICRA '04) 3, 3173–3179.
- Taubin, G., 1991. Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation. *IEEE Trans. Pattern Anal. Machine Intell.* 13 (11), 1115–1138.
- Tong, W.-S., Tang, C.-K., Medioni, G., 2001. First order tensor voting, and application to 3-D scale analysis. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '01), vol. 1.
- Torr, P.H.S., Zisserman, A., 2000. MLESAC: A new robust estimator with application to estimating image geometry. *Comput. Vision and Image Understanding* 78 (1), 138–156. <http://dx.doi.org/10.1006/cviu.1999.0832>.
- Triebel, R., Burgard, W., Dellaert, F., 2005. Using hierarchical EM to extract planes from 3D range scans. Proc. 2005 IEEE Internat. Conf. on Robotics and Automation (ICRA 2005), 4437–4442.
- Turchetto, R., Manduchi, R., 2003. Visual curb localization for autonomous navigation. In: Proc. IEEE/RSJ Internat. Conf. on Intell. Robots and Systems (IROS 2003), vol. 2, pp. 1336–1342.
- Unnikrishnan, R., Hebert, M., 2003. Robust extraction of multiple structures from non-uniformly sampled data. Proc. IEEE/RSJ Internat. Conf. on Intell. Robots and Systems (IROS 2003) 2, 1322–1329.
- Vidal, R., Ma, Y., Sastry, S., 2005. Generalized principal component analysis (GPCA). *IEEE Trans. Pattern Anal. Machine Intell.* 27 (12), 1945–1959. <http://doi.ieeecomputersociety.org/10.1109/TPAMI.2005.244>.