Proceeding of the IEEE
International Conference on Information and Automation
Yinchuan, China, August 2013

Information Theory based Validation for Point-cloud Segmentation aided by Tensor Voting

Ming Liu[1,2] and Roland Siegwart [1]

[1]Autonomous Systems Lab, ETH Zurich, Switzerland
[2]The Hong Kong University of Science & Technology, Hong Kong

***Abstract*—Segmentation of point-cloud is still a challenging problem, regarding observation noise and various constraints defined by applications. These difficulties do not concede to its necessity for almost all kinds of modeling approaches using point-cloud. However, the criteria to justify the quality of a clustering result are not much studied. In this paper, we first propose a point-cloud segmentation algorithm using adapted k-means to cluster normal vectors obtained from tensor voting. Then we concentrate on how to use a non-parametrical criterion to validate the clustering results, which is an approximation of the information introduced by the clustering process. Compared with other approaches, we use noisy point-cloud obtained from moving laser range finders directly, instead of reconstruction of 3d grid-cells or meshing. Moreover, the criterion does not rely on the assumption of distributions of points. We show the distinguishable characteristics using the proposed criteria, as well as the better performance of the novel clustering algorithm against other approaches.**

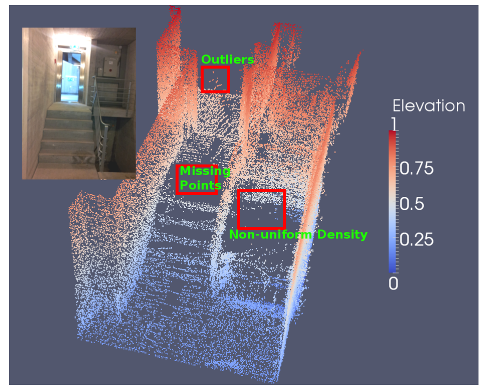


Fig. 1. Clip of a typical point-cloud and common observation noises.

I. Introduction

SEGMENTATION of point-cloud is an important base for several mobile robotic applications, such as topological mapping, semantical reasoning and scene reconstruction. However, it is in general considered as a difficult problem, especially for the applications using raw data points, because of the following major reasons:

- *Unreliable point-cloud observations*: the unreliability is multi-fold. In figure 1, we show a cropped part of point-cloud observed from a doorway environment. Outliers, shadow points, non-uniform distributions, missing points etc. can be observed all over the place. These unreliable observations make proper data-modelings subtle.
- *Application constraints*: since segmentation results of point-cloud are used differently for applications, in order to fit various requirements or constraints, several compromises may be demanded. E.g. the cardinality of clustering sometimes relies on human supervisions.
- *Computational complexity*: we could see from figure 1 that the number of points are usually large. Though sub-sampling is usually performed, the computational cost is still the bottleneck for most applications.

Respecting with these three major difficulties, we tackle the segmentation problem based on raw point-cloud using tensor voting, since it has been proved can help estimate the missing structural information [1]. Regarding the complexity, we use a recently proposed GPU implementation to alleviate the computation [2]. The results are evaluated using a set of datasets for both indoor and outdoor environments [3].

A. Contributions

We address the following aspects in this paper:

- Surface normal estimation using tensor voting. We show its potential for structural extraction.
- Adapted k-means clustering considering both Euclidean distance and surface normal similarity.
- Information theory based evaluation of segmentation results. As an instance, we show the result which uses the proposed criterion for the estimation of the number of clusters (cardinality).

B. Arrangement

The remainder of this paper is arranged as follows. We start with introducing related works by categories. After the overview of the algorithm pipeline in section III, we first discuss The segmentation algorithm and validation methods based on information theory in section IV. In section V, the adapted k-means is introduced as well as the cardinality determination strategy, followed by experiments on real dataset in section VI. At last, we draw conclusions and introduce our vision for future work.

II. Related Work

A. Segmentation of Range Image

Pointcloud is sometimes also referred as “range image”. Usually these two terminologies are not much distinguished.

This work was supported by the EU FP7 project NIFTi (contract # 247870).
liu.ming.prc@gmail.com, rsiegwart@ethz.ch

978-1-4799-1334-3/13/$31.00 ©2013 IEEE

In this paper, we consider range images as the 3d data which can be considered captured by a single scan, namely a grabbed *frame* from a 3D range finder. Point-cloud can be more complex in general. It can be combined with several registered scans. It implies that the structure can not be directly reprojected to a 2D plane without losing information.

Several works regarding range image segmentation have been proposed. These works are based on different features, such as edge, projected as from 2d images [4], [5]. Then computer vision techniques are applied to process the data in 2D. The main drawback of these works is that they rely on almost-clean dense representations of the target models, which is not the case for most robotics applications. Moreover, the back-forth projection between 2d and 3d representation is time consuming. The ideal case is that all scans are with 0-noise and uniformly distributed points. For typical mobile robotic applications, since the range finder is mounted on a moving platform and surveillant the work space from different viewpoints, these problems are however inevitable.

The structural information can be obtained by local parametrical modeling (e.g. surface model [6], [7]), non-parametrical regression [8], and other local embedding techniques [9]. Regarding the propagation, they can be broadcast over field [10], structured grid [11], [12], or over unstructured graph [13] etc. Douillard et al showed a more detailed comparison among common parametrical modeling algorithms [12]. They discussed the metric for segmentation validation should be based on a manually labeled ground truth. We consider the assessment of the clustering result is able to be carried out automatically with a non-parametrical manner.

B. Information Theory

In contrast to semi-supervised algorithms [14], for most unsupervised clustering algorithms, such as spectral clustering [15] or K-means, the target number of clusters of input data is a key issue, namely *cardinality* of clustering. This problem has been extensively studied in [16]. Though several adaptations are proposed for specific applications such as self-tuning spectral clustering [17], this problem persists in general. In order to eliminate the dependency on knowing cardinality as aprior, usually information theory can be adopted, such as using Dirichlet Process based modeling [18]. In our previous work [19], we proposed a segmentation algorithm using Chow-Liu decomposition of the mutual information tree. When local constraints can be defined, the cardinality can be taken as parameter of global entropy maximization. Mutual information has been proved to be an optimal clustering criterion [20] for such setup. When visual features are introduced, several segmentation applications can be found. They cover a range from scene recognition [21], [22] and visual homing based navigation [23], [24] etc.

C. Normal based Segmentation

Surface normal is a local consistent feature. Therefore it is widely used for point-cloud analysis. Regarding segmentation, one early work by Pulli et al [25] aims at segmenting range images into homogeneous regions, by decomposing x- and y-components of the normal vectors. It assumes perfect dense point clouds and the resulting algorithm only deal with segmentation in 2.5D. Normal estimation based on local constrained least square modeling [26] was also studied, but it has much higher complexity for the regression process. Holz et al proposed a segmentation algorithm by an initial segmentation in normal space, then refines in distance space, without specified clustering algorithm [27]. Teutsch et al presented a clustering algorithm for subset segmentation [28], which targets at segmentation of point clouds without plane-assumption. [29] introduced an incremental way to model different clusters by using both angular and distance constraints. In this work, we fuse the distance in normal direction space into a compact distance function, where the normal difference is used as a weighted component.

D. Tensor voting

Tensor voting [1] is originated in computer vision. It has been extended to several applications related to segmentation [30], [31]. Several works have been proposed on structure extraction of point-cloud using Tensor voting as well [32], [33]. We consider it is one of the most important algorithms for structural analysis, because it is extraordinary performance in its tolerance to noise and missing data, its consistency for local information and intuitive extraction of evidence saliency etc. Nevertheless, the computational cost of Tensor voting is high. Thanks to the CUDA based tensor voting open-source library, which we recently proposed [2], the algorithm is able to launch in near real-time for reasonable data size. As a result, both surface normals and geometrical saliency can be extracted.

III. Overview

The pipeline of the proposed algorithm is shown as figure 2. We start with a sampling process of the raw data based on local

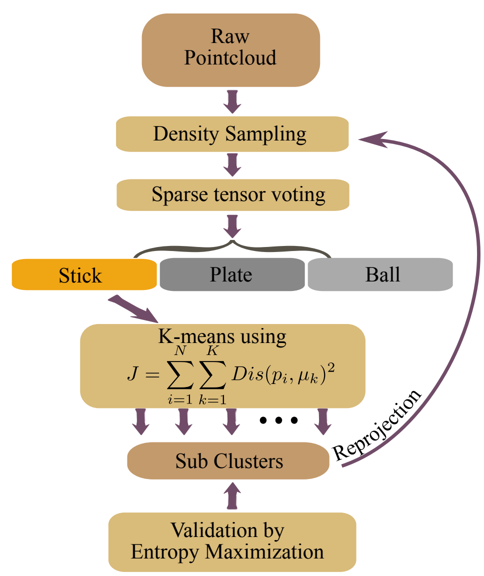


Fig. 2. The pipeline of the proposed algorithm

density. This step is critical for tensor voting, as discussed in [33], because the voting result is a collective sum of all voters surrounding a certain point. The density sampling process, as a standard process, will greatly help to reduce the influence of the non-uniform distribution of points. The samples are then processed by sparse tensor voting. After that, the resulting *stick* components are used to describe the similarity in normal space among all the points. We use adapted k-means to perform clustering over the points by a redesigned energy function J, which is discussed in section V-A. In order to validate the clustering results, as well as select the proper cardinality of clustering, a criterion based information theory is evaluated. The details can be found in the following sections.

IV. Segmentation Quality Assessment by Information Theory

A. Information based quality assessment

We consider the segmentation problem as distributional clustering. It means that each point is described as a distribution. The relative entropy then becomes a natural measure of the distance between distributions. Each distribution is modeled as two parts: the distance in the normal space to the cluster mean normal and local embedding relations to kNN in its own cluster. For a cluster b, these two models are represented by M^b_{normal} and M^b_{knn} respectively.

Given such a choice, mutual information has been proved to be an optimal clustering criterion [20]. The clustering of point-cloud $P = \{p_1, p_2, \ldots, p_N\}$ into K clusters is a function:

$$C : P \to \{1, \ldots, K\}$$

K is named as the cardinality of clustering.

By introducing clustering function C, the data points are better organized. Therefore, we want to maximize the mutual information between raw point-cloud and the clustering, namely $I(P, C)$. Based on information theory, the mutual information can be further representation by

$$I(P, C) = H(P) - H(P \mid C) \tag{1}$$

In general, the goal of clustering is to maximize the divergence between clusters and minimize the "dynamics" within the same cluster. In this case, $H(P)$ is determined with the raw data, therefore independent on any specific clustering. The target problem is thus converted as an optimization problem, respecting with the minimization of a score in the following form:

$$S_{MI}(C) = H(P \mid C) = \sum_{k=1}^{K} \frac{n_k}{N} H(P \mid C = k) \tag{2}$$

where n_k is the number of points assigned to cluster k, N is the total number of points. Score S_{MI} describes the average entropy of the data points from clustering function C. For an optimal clustering C, the score will be minimized, resulting in maximized mutual information $I(P, C)$.

For each point p^k in cluster k, the likelihood of occurrence is:

$$p(p^k \mid C = k) \propto \sum_{M^k_{normal}}^{K} \sum_{M^k_{knn}}^{K} p(p^k \mid M^k_{normal}, M^k_{knn}) \\ p(M^k_{normal} \mid C = k) p(M^k_{knn} \mid C = k) \tag{3}$$

Here we assume the two data models M^b_{normal}, M^b_{knn} are independent conditional on knowing clustering C. Thus the intra-cluster entropy:

$$H(P \mid C = k) = -\sum_{p^k}^{n_k} p(p^k \mid C = k) \log p(p^k \mid C = k) \tag{4}$$

We need to see that there are several problems hinder the performance of the clustering using the score function (2), considering the formation of (3). Firstly, the joint probability $p(p^k, M^k_{normal}, M^b_{knn})$ is hard to formulate in a parametrical way, which makes it hard to be estimated, e.g. by Maximize-Likelihood Estimation. Besides, the derivation of (4) is not possible to calculate in a close-form due to the first issue. Last but not least, the cardinality of clustering K is unknown. In the worst case, finding an optimal solution of unknown cardinality of clustering is NP hard. Therefore, a proper approximation of the joint distribution of data and models is required, which is preferable to be independent of data distribution. We use non-parametrical methods to approximate it, as introduced in the next subsection.

B. Non-parametrical approximation of intra-cluster Entropy

Faivishevsky et al proposes a non-parametrical information clustering (NIC) estimator using MeanNN differential entropy [8], which leads to an inter-cluster entropy estimation that can be summarized by:

$$H(X \mid C = j) \approx \frac{d}{n_j(n_j - 1)} \sum_{i \neq l \mid c_i = c_l = j} \log ||x_i - x_l|| \tag{5}$$

where $|| \cdot ||$ is the Euclidean distance between two datapoints x_i and x_l (in $\mathbf{R}^d$). We redesign this distance measure by including the similarity in surface normal space as well in order to enhance robustness and distinct the local planes. We use the following measure in Euclidean-Normal space, for two points p_i and p_l:

$$Dis(p_i, p_l) = D_{metric}(p_i, p_l) + \lambda D_{angle}(p_i, p_l) \tag{6}$$

$$D_{metric}(p_i, p_l) = ||p_i - p_l|| \tag{7}$$

$$D_{angle}(p_i, p_l) = 1 - \frac{\overrightarrow{p_i} \cdot \overrightarrow{p_l}}{|\overrightarrow{p_i}||\overrightarrow{p_l}|} \tag{8}$$

where $\overrightarrow{p_i}$ and $\overrightarrow{p_l}$ are normal vectors for points p_i and p_l respectively. Parameter λ is used to adjust the weight of angular difference. We can infer that it is a parameter that converts the difference from normal space to metric space, since the D_{angle} is within the range $(0, 1)$.

By combining (2 5 6), we can rewrite equation 2 in the NIC form as [1]:

$$S_{NIC}(C) = \frac{1}{N}\sum_{k}\frac{3}{n_k - 1}\sum_{i\neq l|c_i=c_l=j}\log Dis(x_i - x_l) \quad (9)$$

The NIC form score is to evaluate the quality of the clustering. In [8], the calculation and naive clustering assignment is executed in a non-heuristic way, where the reassignments of cluster identity is random. It leads to very high computational complexity of $O(n^2)$. Considering the large number of points in a typical point-cloud, the calculation will be extremely slow, let alone the issue of unknown cardinality. In this paper, instead of searching for optimal solution derived by (9), we use (9) to assess the quality of clustering results. Compared with other information based criterion, such as Bayesian information criterion (BIC):

$$BIC(C|X) = -2\cdot L(X|C) - \varphi\cdot\log n \quad (10)$$

where φ is the number of free parameters, $L(\cdot)$ is the log-likelihood, the NIC form is calculated from a pair-wise distance matrix directly. However, the BIC criterion is not directly applicable in this case, since the distribution function is not explicit. After all, we see the following two benefits of (9). On one hand, it greatly facilitates the programming and calculation; on the other hand, the non-parametrical form enables the assessment of the quality to be independent of model assumptions.

Another widely cited criteria is the so-called Clustering Evaluation Function (CEF) [34] based on quadratic form of Renyi's entropy measure [35].

$$S_{CEF}(C) = \frac{1}{2N^2}\sum_{i=1}\sum_{j=1}M(p_i,p_j)G(p_i - p_j, 2\sigma^2) \quad (11)$$

where $G(x, 2\sigma^2)$ is the value calculated by a Gaussian probability density function kernel $N(0,\sigma)$. Parameter $M(p_i,p_j)$ is an indicator which is 1 when p_i and p_j are labelled differently, otherwise 0. We will further compare these two criteria in the next section, based on the experiment results.

V. Adapted K-means Clustering

A. K-means using Adapted distance

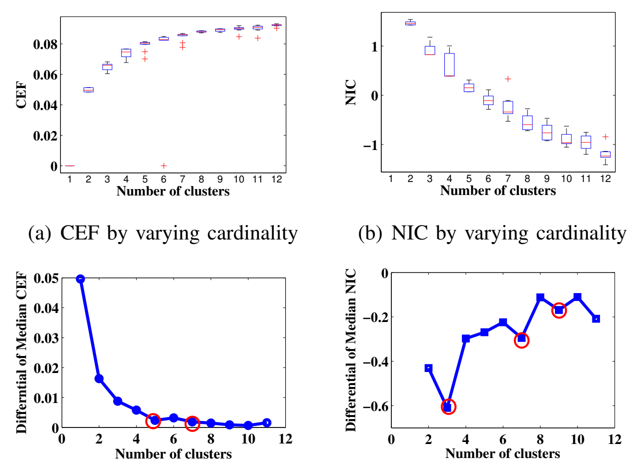
The non-uniform distribution and missing points from point-cloud imply that a standard k-means clustering can not lead to correct result, which is also shown in the comparison given in the next subsection. Moreover, we could not use parametric model based clustering either, e.g. using Expectation Maximization (EM) to tune parameters afterwards, because any two arbitrary points are independently distributed. Here we propose a K-means algorithm using the adapted distance function proposed in 6. Inspired by soft K-means [36], where an entropy term is used to encourage the responsibility of each point as equalized as possible, we adopt information theory based criteria to indicate the local similarity within each cluster. The cost energy function is formed as follows:

$$J = \sum_{i=1}^{N}\sum_{k=1}^{K} Dis(p_i, \mu_k)^2 \quad (12)$$

Applying the k-means routine using the cost energy function (12), the complexity of algorithm is reduced to $O(NK)$, comparing with naive reassignments $O(N^2)$.

B. Determination of cardinality

We recall that the determination of cardinality K is generally required. As discussed previously, we use the NIC score to find the optimal. Meanwhile, CEF is also [34] evaluated as comparison. By varying the size of K, different scores after convergence can be obtained as shown in figure 3. It shows the statistics of 20 tests for each cardinality candidate. Figure 3(a) and (b) show box-plots over all candidates using CEF and NIC respectively. The computational complexities of both criterion are $O(N^2)$. The changes of both criteria are monotonic along the variation of the number of clusters. However the difference between two neighbouring choices is of interest. In order to eliminate the influence of singularities, we first extract the median values of the two groups of tests. After that, we calculate the adjacent differences of the median values, as shown in figure 3(c) and (d). The red circles marks the potential number of clusters. They highlight the valley of major slope transitions. We could see that the CEF criterion is less intuitive than NIC. Moreover, NIC has less requirement from data, since it is independent of the real data distribution. On the contrary, the Renyi's Entropy relies on statistics based on local Gaussian kernels [37].



(a) CEF by varying cardinality (b) NIC by varying cardinality

(c) CEF Variation (d) NIC Variation

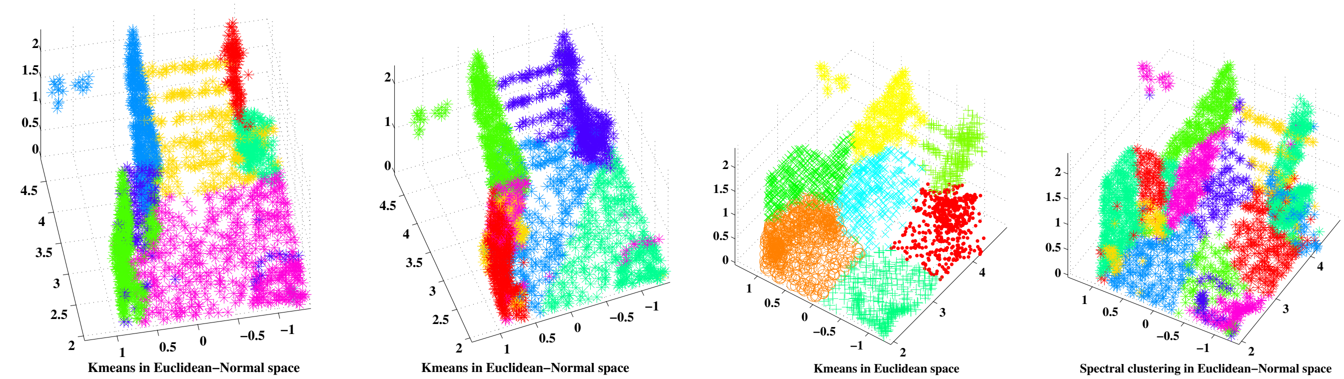
Fig. 3. Determination of cardinality of clustering. Red circles in (c) (d) marks the detected salient changes, which indicates the potential number of clusters. We could see that NIC criteria is more distinguishable in cardinality detection.

C. Clustering results on a typical indoor scan

The clustering results for the dataset depicted in figure 1 is shown in figure 4(a) and (b) using K=7 (derived from figure 3),

[1]In the original paper [8], the $\frac{1}{N}$ nominator was missing.

Result using proposed algorithm	Result of bad initialization	Standard k-means	Spectral clustering



(a)	(b)	(c)	(d)
$S_{NIC} = -0.4541$	$S_{NIC} = -0.2110$	$S_{NIC} = -0.0292$	$S_{NIC} = -0.0175$
$S_{CEF} = 0.0922$	$S_{CEF} = 0.0772$	$S_{CEF} = 0.0865$	$S_{CEF} = 0.0871$

Fig. 4. Clustering results using different algorithms on the same structured dataset (after density sample: 2588 points). A lower S_{NIC}, or higher S_{CEF} indicates better result, according to definitions. We could see that S_{NIC} criterion is more informative as a assessment of the segmentation quality.

as a typical choice based on the result in 3(d). The different converged results are caused by random initial states of k-means. Though it can be optimized by more sophisticated initial state selection algorithms, such as one seed per normal cluster etc, this problem is not further investigated in this paper. Nevertheless, the distinguished scores S_{NIC} indeed show the quality difference, regarding rationality of clustering. We compare the proposed algorithm with two other methods. First, standard k-means is executed based on Euclidean distance. The result is shown in figure 4(c). The clusters are equally distributed. This result is less helpful for further analysis, such as structure reasoning. Besides, we perform spectral cluster by using the same distance function (6). From the result depicted in figure 4(d), we could see that some part of the point cloud, such as side-walls are corrected clustered. However, the result is not satisfied in terms of compactness and meaningful segments. Moreover, the computational cost is much higher than basic k-means, since spectral clustering conducts with an inversion of the $N \times N$ distance matrix. By comparing the two criteria, we could observe that S_{NIC} is more distinctive the S_{CEF}. Especially for the non-preferable results in figure 4(b,c,d), the quality differences represented by S_{NIC} is more rational.

VI. Further Experiments and Discussion

A. Result on a typical outdoor dataset

A result on outdoor datasets can be found in figure 5. It is a dataset with a tree on the ground. This dataset is an extreme case which contradict with several assumptions we have in this paper. Especially the normal directions from the tree crown is very badly distributed. Therefore, for other approaches that based only on normal clustering will not work. We see that the distance function proposed in (6) combines the similarities in both Euclidean space and surface normal space, results in rational results. By cardinality analysis, we could see that the rational clustering can be performed with cardinality 2, 4, 7 or 9. We show the result of 2 and 4 clusters in figure 5(c) and (d), which segments the dataset into parts of ground, tree branch and crown. The S_{NIC}'s for these results are also provided as benchmark for further studies.

VII. Conclusion and Future Work

In this paper, we have studied several criteria for assessing the quality of clustering problem in terms of modeling of point-clouds. Meanwhile, we proposed a k-mean inspired clustering algorithm using a distance function fusing the surface normal and Euclidean distances. The surface normals are obtained by sparse tensor voting algorithm, which allows adaptive structural analysis depending on the kernel size. The clustering results are validated by non-parametric criteria. The results show that it can be easily applied to point-cloud segmentation problem, and obtain rational results. Further automated point-cloud modeling algorithms can be envisaged.

References

[1] G. Medioni, M. Lee, and C. Tang, *A computational framework for segmentation and grouping*. Elsevier Science, 2000, vol. 1.

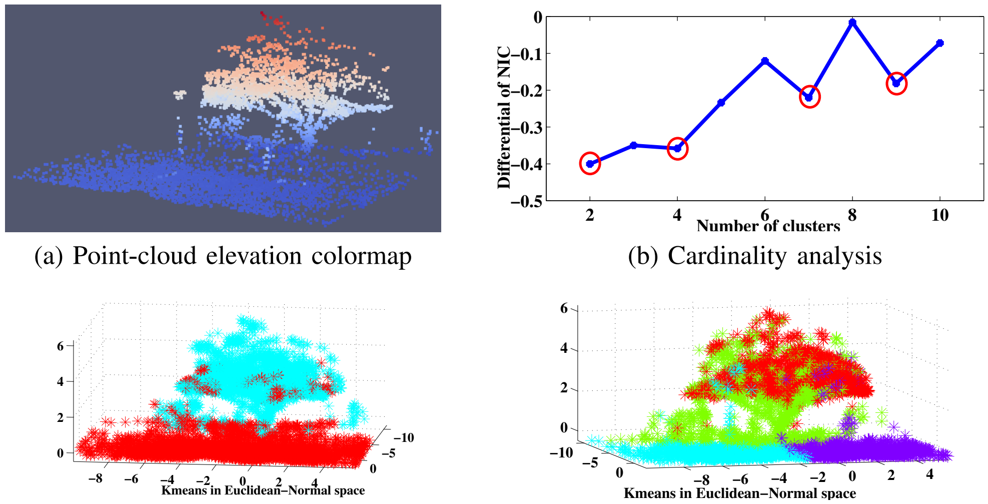
[2] M. Liu, F. Pomerleau, F. Colas, and R. Siegwart, "Normal Estimation for Pointcloud using GPU based Sparse Tensor Voting," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics*. IEEE, 2012.

[3] Autonomous Systems Lab, ETH Zürich, "ASL Datasets Repository." [Online]. Available: {http://http://projects.asl.ethz.ch/datasets/doku.php}

[4] Y. Alshawabkeh, N. Haala, and D. Fritsch, "Range image segmentation using the numerical description of the mean curvature values," in *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. ISPRS Congress*, 2008, p. 533.

[5] H. Iddamsetty, "Segmentation of range images for modeling of large outdoor scenes," 2003.

[6] A. Leonardis, A. Gupta, and R. Bajcsy, "Segmentation of range images as the search for geometric parametric models," *International Journal of Computer Vision*, vol. 14, no. 3, pp. 253–277, 1995.



(a) Point-cloud elevation colormap (b) Cardinality analysis

(c) Result $S_{NIC} = 4.806$ (d) Result $S_{NIC} = 4.467$

Fig. 5. Results on an outdoor dataset with a tree (after density sample: 6000 points)

[7] G. Yu, M. Grossberg, G. Wolberg, and I. Stamos, "Think globally, cluster locally: A unified framework for range segmentation," in *International Symposium on 3D Data Processing, Visualization and Transmission*, 2008.

[8] L. Faivishevsky and J. Goldberger, "A nonparametric information theoretic clustering algorithm," in *ICML*. Citeseer, 2010, pp. 351–358.

[9] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.

[10] M. Johnson-Roberson, J. Bohg, M. Björkman, and D. Kragic, "Attention based active 3d point cloud segmentation," *IROS 2010*, 2010.

[11] H. Woo, E. Kang, S. Wang, and K. Lee, "A new segmentation method for point cloud data," *International Journal of Machine Tools and Manufacture*, vol. 42, no. 2, pp. 167–178, 2002.

[12] B. Douillard, J. Underwood, N. Kuntz, V. Vlaskine, A. Quadros, P. Morton, and A. Frenkel, "On the segmentation of 3d lidar point clouds," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2798–2805.

[13] A. Golovinskiy and T. Funkhouser, "Min-cut based segmentation of point clouds," in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*. Ieee, 2009, pp. 39–46.

[14] M. Liu, F. Colas, F. Pomerleau, and R. Siegwart, "A Markov semi-supervised clustering approach and its application in topological map extraction," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012.(IROS 2012)*, 2012.

[15] T. Hastie, R. Tibshirani, and J. J. H. Friedman, *The elements of statistical learning*. Springer, 2009.

[16] H. Xiong, J. Wu, and J. Chen, "K-means clustering versus validation measures: a data-distribution perspective," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 39, no. 2, pp. 318–331, 2009.

[17] L. Zelnik-Manor and P. Perona, "Self-tuning spectral clustering," *Advances in neural information processing systems*, vol. 17, no. 1601-1608, p. 16, 2004.

[18] M. Liu, L. Wang, and R. Siegwart, "DP-Fusion: A generic framework for online multi sensor recognition," in *Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2012.

[19] M. Liu, F. Colas, and R. Siegwart, "Regional topological segmentation based on mutual information graphs," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011.

[20] A. Banerjee, S. Merugu, I. Dhillon, and J. Ghosh, "Clustering with bregman divergences," *The Journal of Machine Learning Research*, vol. 6, pp. 1705–1749, 2005.

[21] M. Liu, R. S. Davide Scaramuzza, Cédric Pradalier, and Q. Chen, "Scene recognition with omnidirectional vision for topological map using lightweight adaptive descriptors," in *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*, 2009.

[22] M. Liu and R. Siegwart, "DP-FACT: Towards topological mapping and scene recognition with color for omnidirectional camera," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA 2012)*, 2012.

[23] M. Liu, C. Pradalier, F. Pomerleau, and R. Siegwart, "The role of homing in visual topological navigation," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012.(IROS 2012)*, 2012.

[24] ——, "Scale-only visual homing from an omnidirectional camera," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, may 2012, pp. 3944 –3949.

[25] K. Pulli and M. Pietikäinen, "Range image segmentation based on decomposition of surface normals," in *Proceedings of the Scandinavian conference on image analysis*, vol. 2, 1993, pp. 893–893.

[26] E. Castillo and H. Zhao, "Point cloud segmentation via constrained nonlinear least squares surface normal estimates," Technical Report CAM09-104, Computational and Applied Mathematics Department, University of California Los Angeles, Tech. Rep., 2009.

[27] D. Holz, S. Holzer, R. Rusu, and S. Behnke, "Real-time plane segmentation using rgb-d cameras⋆," in *Proc. of the 15th RoboCup International Symposium*, 2011.

[28] C. Teutsch, E. Trostmann, and D. Berndt, "A parallel point cloud clustering algorithm for subset segmentation and outlier detection," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 8085, 2011, p. 8.

[29] K. Klasing, D. Wollherr, and M. Buss, "Realtime segmentation of range data using continuous nearest neighbors," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 2431–2436.

[30] Y. Dumortier, I. Herlin, and A. Ducrot, "4-d tensor voting motion segmentation for obstacle detection in autonomous guided vehicle," in *Intelligent Vehicles Symposium, 2008 IEEE*. IEEE, 2008, pp. 379–384.

[31] J. Jia and C. Tang, "Inference of segmented color and texture description by tensor voting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 6, pp. 771–786, 2004.

[32] H. Schuster, "Segmentation of lidar data using the tensor voting framework," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 35, pp. 1073–1078, 2004.

[33] B. King, "Range data analysis by free-space modeling and tensor voting," Ph.D. dissertation, RENSSELAER POLYTECHNIC INSTITUTE, 2009.

[34] E. Gokcay and J. Principe, "Information theoretic clustering," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 2, pp. 158–171, 2002.

[35] A. RRNYI, "On measures of entropy and information," in *Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1961, pp. 547–561.

[36] J. Kim, K. Shim, and S. Choi, "Soft geodesic kernel k-means," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 2. IEEE, 2007, pp. II–429.

[37] J. Principe, D. Xu, and J. Fisher, "Information theoretic learning," *Unsupervised adaptive filtering*, vol. 1, pp. 265–319, 2000.