# Object extraction based on 3d-segmentation of LiDAR data by combining mean shift with normalized cuts: two examples from urban areas

Wei Yao[1], Stefan Hinz[1] and Uwe Stilla[2]
[1]Remote Sensing Technology, [2]Photogrammetry and Remote Sensing
Technische Universitaet Muenchen, 80333, Munich, Germany
{Wei.Yao, Stefan.Hinz@bv.tum.de}; Stilla@tum.de

***Abstract*— In this work, we have looked into the problem of urban analysis using airborne LiDAR data based on the strategy of classification by segmentation. Segmentation is a key and hard step in the processing of 3D point clouds, which is not perfectly solved in view of different applications. A new 3d segmentation method incorporating the advantages of nonparametric and spectral graph clustering is presented here to facilitate the task of object extraction in urban areas. This integrated method features local detection of arbitrary modes and globally optimized organization of segments concurrently, thereby making it particularly appropriate for partitioning raw airborne LiDAR data of urban areas into segments approximating semantic entities. Two examples in urban areas — flyover and vehicle are chosen as interest objects to be extracted by a classification–based step. The approach has been tested on LiDAR data of dense urban areas, and the results that are obtained have been compared with manual counts and showed us the efficiency and reliability of the strategy.**

## I. INTRODUCTION

Airborne Laserscanning (ALS) data has become an important source for object detection and reconstruction for various applications such as urban and vegetation analysis [1] [2]. However, since the raw ALS data are unstructured 3d point clouds that are unevenly sampled throughout the real world depending on acquisition conditions, a key challenge in the automatic processing of laser data is segmentation. It refers to the partition of a set of measurements in the 3d object space (point cloud) into smaller and coherent subsets that could later lead to a proper recognition of object-classes. In this paper we believe that segmentation plays a more important role towards handling ALS data semantically, and can roughly be divided into two strategies:

**Model-driven**: This strategy assumes models for describing segments by parametric geometry, e.g. planes or low-order polynomials surfaces), the segmentation and reconstruction of the primitives are typically performed simultaneously and usually work on gridded range image data, namely 2.5d point cloud with a fixed, rectangular 2d grid topology [3][4]. Recently, approaches for shape detection running directly on the 3d unstructured point cloud via robust fitting [5] or parameter-space analysis [6] have emerged and can also be assigned to this class.

**Data-driven**: This strategy is able to work on both the original 3d points and the gridded data. It is motivated by the recognition that points that constitute a segment tend to cluster if represented by adequate features. The feature-based clustering was chosen mainly due to the generality and flexibility it offers in accommodating spatial relations among points, observed attributes, and derived attributes [7]. The features often considered are based on measures derived from localized regions or points, e.g. height texture, curvature, normal vector and intensity etc. A newly developed method [8] showed that the classification can also be done in object-based wise.

The work described in this paper has been motivated by subsequent considerations: First, range image segmentation has inherent limitations because it requires that the unstructured point cloud is rasterized or interpolated into a regular grid. 3-dimensional internal structures present in the original data (e.g. power cables situated above each other, jutties, and several layers of branches in a tree) are squashed into 2.5D and thus lost. Second, a severe obstacle to model-based segmentation is the fact that a complex real-world scene, e.g., a city area, will comprise many different kinds of objects or clutters, and ALS data are subject to a diversity of topology, density and point distribution. No single geometric primitive is likely to describe all of objects sufficiently well. Although semi-parametric approaches to clustering, which assume a generic geometric or statistical model for the individual clusters, e.g. expectation maximization, were used, for real-world problems any model assumption is not preferred. In addition, most of clustering methods currently adopted (e.g. K-means) demands that the number of clusters should at least be apparent from prior knowledge, but the determination of this parameter is not a trivial problem. Third, so far, most approaches devoted to urban object extraction from ALS data are application-specific and lack of generality[9][10][11],121314 which means that they focused on detecting and highlighting single object class from urban areas rather than firstly distinguishing between various ones. Therefore, model-based and context-guided strategies were often used for these cases. Our question has

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

emerged as follows: can we do the segmentation and extraction of urban objects from laser data simultaneously and achieve the semantic labeling of urban objects from ALS data based on a special segmentation framework towards LiDAR data?

This paper is to introduce a new general segmentation strategy for extracting objects from ALS data of urban areas based on nonparametric and spectral graph clustering,

- which is able to adaptively cope with the undesirable diversity of laser data properties, such as topology, quality, density, and multi-modality.
- which can run in the feature space spanned by feature vectors that are not biased by local neighborhood selection, namely working genuinely on 3D point cloud..
- which can by incorporating global optimization criterions provide a reasonable partition of ALS data towards the facilitation of object recognition..

## II. General Strategy

Our research in this paper is to introduce a new segmentation strategy for extracting objects from ALS data of urban (Fig.1). It deals with semantic inference in the 3d point cloud data, which has proved to be more complicated than in image data due to the inherent characteristics. We take flyover and vehicle in urban areas as two example objects to be extracted, because they were rarely handled or regarded as inelegantly solved in previous researches; moreover, they are objects at two totally different scales.

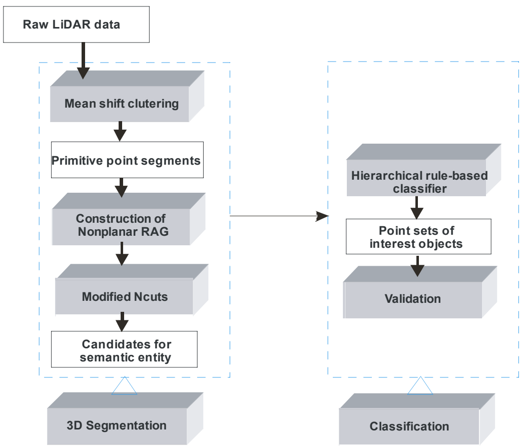



Figure 1. Flow chart of the general strategy

A classification by segmentation strategy is applied to ALS data in order to extract the two object classes simultaneously. The ALS data is firstly subjected to the segmentation process that is consisted of nonparametric clustering tool – mean shift (MS). The obtained result is usually not able to provide significative regions to represent distinct natural and man-made characteristics of the complex scene, despite that MS can do a genuine clustering directly on 3d point cloud and discover various geometric modes in it. This deficiency of MS segmentation is especially evident in terms of the separation of objects at completely different scale levels, owing to the uniquely adjustable parameter — bandwidth *h*. Therefore, the initial resulted point segments have to be handled in a second step considering the global optimization criterions, to generate more consistent subsets of laser points. Since the number of point segments yielded by MS is larger than the number of typical scene-objects, a modified normalized-cuts (Ncuts) within the scope of perception grouping is proposed in this step. It adapts ordinary Ncuts to the task of grouping 3d point segments by constructing non-planar adjacency graph with two types of edge connecting nodes. Finally, the raw laser data is spatially separated into point clusters that potentially correspond to semantic object entities, and then classification is performed to evaluate the point clusters to extract the flyover and vehicle in urban areas based on shape and positional criterions.

## III. Segmentation

### A. *Initial partition by mean shift*

Mean shift is an extremely versatile tool for feature-space clustering. It has its roots in non-parametric density estimation. A remarkable feature is its ability to move the data points towards their respective modes of the empirical distribution without having to estimate or evaluate the underlying probability density model itself. So far, MS has been successfully applied to image processing fields, such as the segmentation of images.

Melzer [12] has initialized the work of using MS to explore urban ALS data and gave us a very promising prospect of it concerning scene segmentation. Thanks to its non-parametric and mode-seeking mechanism, we can obtain a genuine clustering of 3d point cloud without enforcing any possibly inadequate model assumptions. Therefore, MS is believed to have great potential with respect to flexibility and reliability when deal with laser data of complex urban areas. For our case, a partition and description of laser data primitives towards object extraction is constructed via MS.

Given $n$ laser data points $\mathbf{x}_i$ , $i=1,\ldots,n$ in the 3-dimensional object space $R^3$, we consider radially symmetric kernels satisfying $K(\mathbf{x}) = c_{k,d}\, k(\|\mathbf{x}\|^2)$, where $c_{k,d} > 0$ is chosen such that $K(\mathbf{x})$ integrate to one. $k(x)$ is referred to as the kernel profile and defined only for $x > 0$. Now given the function $g(x) = -k'(x)$ for profile, the kernel $G(\mathbf{x})$ is defined as $G(\mathbf{x}) = c_{g,d}\, g(\|\mathbf{x}\|^2)$, MS vector is defined as

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^{n} \mathbf{x}_i\, g\left(\left\|(\mathbf{x}-\mathbf{x}_i)/h\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|(\mathbf{x}-\mathbf{x}_i)/h\right\|^2\right)} - x , \quad (1)$$

where x is the center of the kernel (window), and $h$ is a bandwidth parameter. MS vector computed with kernel $G$ is proportional to the density gradient estimate obtained with kernel and thus points toward the direction of maximum increase in the density. Therefore, MS executed recursively is able to converge at a nearby point where the density estimate has zero gradient [13]. The modes of feature space are usually located among the region in which the majority of the points

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

reside, equaling to zeros of the gradient, and consequently they are such stationary points of the density where MS can define a path leading to. MS is an adaptive gradient ascent method; the center of the kernel $K$ can be updated iteratively by

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^{n} \mathbf{x}_i g\left(\left\| (\mathbf{y}_j - \mathbf{x}_i)/h \right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\| (\mathbf{y}_j - \mathbf{x}_i)/h \right\|^2\right)}, \quad j = 1, 2, \ldots \qquad (2)$$

where $\mathbf{y}_1$ is the centre of the initial kernel. Once $\mathbf{y}_j$ gets sufficiently close to a mode of the estimated density, it converges to it. The set of all locations that converge to the same mode defines the cluster of that model.

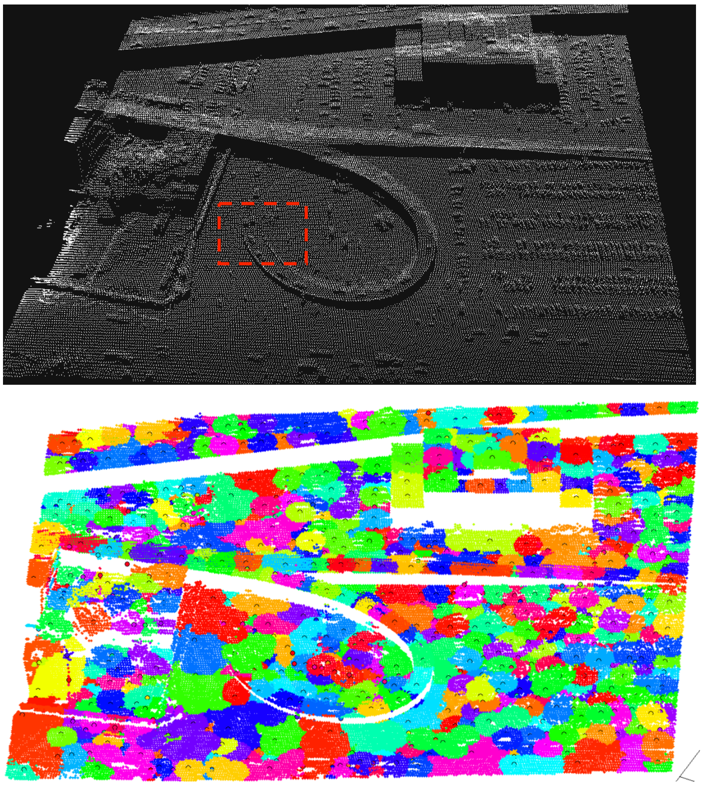


Figure.2. Top: original ALS data, bottom: segmented result using MS with mode centre marked by black circles

Based on the above analysis, MS can easily be adapted to the clustering task for laser point clouds intended for finding various geometric modes, when directly assuming the 3d Euclidean space of laser points as feature space. Since the raw laser data feature unstructured data format, it is intuitive for MS to directly obtain a genuine 3d clustering on it without any preprocessing. It can be assumed that the density modes found in laser data tend to emerge about local distinct structures in the 3d geometric object space, to form a weak primitive description for the scene. One example is demonstrated in Fig.2. Only one parameter that has to be specified in advance for MS procedure are the kernel width $h$ (optionally, the minimum number $M$ of points in a region may also be defined), which can be perceived as the scale factor for distinct structure. The kernel can be distinguished between an isotropic and anisotropic kernel based on whether the width for all dimensions is same or not. For our application, an anisotropic Gaussian kernel with $h_v$ for vertical dimension and $h_p$ for two horizontal dimensions is preferred. In addition, k-d tree data structure was used for laser data to accelerate multi-dimensional range search. It is a key factor in applying MS to ALS data which often contain a large number of points.

### B. *Global grouping by normalized cuts*

However, as depicted in Fig.2, it is difficult to partition a 3d urban scene into significant regions to adequately represent distinct and meaningful terrain or man-made characteristics, relying only on the MS segmentation. It is lack of scale-adaptivity. The main reason lies in the fact that MS algorithm is an unsupervised clustering-based segmentation method, where the number and the shape of the data cluster are unknown a-priori. Moreover, the number of segmented regions is mainly determined by the predefined parameters: $M$ and $h$. MS often results in over-segmentation and produces a large number of small but quasi-homogenous sets of points. These point segments are merely meant to approximate the scene (sub)primitives. Consequently, the ideal of spectral graph partitioning is introduced in the sense of perceptual organization, attempting to refine these point segments to form semantic objects. Spectral graph partitioning methods have been successfully applied to many areas in computer vision. Here we use one of these techniques, namely, the normalized cuts method [14], for grouping point segments.

A graph-partitioning method attempts to organize nodes into groups such that the intergroup similarity is low. Given a graph ***G* = (*V*,*E*,*W*)**, where $V$ is the set of nodes, and $E$ is the set of edges connecting the nodes. A pair of nodes $u$ and $v$ v is connected by an edge and is weighted by $w(u,v) = w(v,u) \geq 0$ to measure the dissimilarity between them. $W$ is an edge affinity matrix with $w(u,v)$ as its $(u,v)$ element. The graph can be partitioned into two disjoint sets $A$ and $B = V - A$ by removing the edges connecting the two parts. This problem can be mathematically formulated as graph-cuts by minimizing the *cut* value.

$$cut(A,B) = \sum_{u \in A, v \in B} w(u,v) \qquad (3)$$

Normalized cuts is a new graph-cuts method and superior to other spectral graph partitioning methods due to jointly considering the intragroup similarity within the total measure[15]:

$$Ncut(A,B) = \frac{cut(A,B)}{assoc(A,V)} + \frac{cut(A,B)}{assoc(B,V)} \qquad (4)$$

where $assoc(A,V)$ denotes the total connection from nodes in $A$ to all nodes in the graph, and $assoc(B,V)$ is similarly defined.

In our case, the Ncuts concerns the ALS data in a larger scale and is applied to the results obtained by MS instead of the single points. The point segments produced by MS segmentation can be represented by a Region Adjacency Graph (RAG) that incorporates the topological information of the laser data and region connectivity. Two factors that play a key role in obtaining the good performance lie in the construction

978-1-4244-3461-9/09/$25.00 ©2009 IEEE

of RAG for nodes by analyzing and determining the connectivity between every 3d point segments, and the definition of the dissimilarity measure between neighboring nodes by incorporating most application-relevant and distinguishing features.

A key novelty of the proposed approach for using the Ncuts for grouping the 3d point segments is to introduce a non-planar RAG to describe 3D topological relations and spatial connectivity between every point segments (Fig.3), where the graph edge is defined into two types:

given the premise: $\min\left\{\|dist(R_i,R_j)\|^2\right\} < DistThresh$

●horizontal edge (***HE***),

$$\text{if } \arccos\left(\frac{N_V\{R_i\}\cdot N_V\{R_j\}}{|N_V\{R_i\}|\,|N_V\{R_j\}|}\right) \approx 0^\circ \wedge N_V\{R_i\} \perp C_{Rij}$$

●spatial edge (***SE***),

$$\text{if } \arccos\left(\frac{N_V\{R_i\}\cdot N_V\{R_j\}}{|N_V\{R_i\}|\,|N_V\{R_j\}|}\right) \gg 0^\circ \vee N_V\{R_i\}\ \text{not} \perp C_{Rij}$$

where assume that the laser data was by MS segmented into $n$ regions $R_i$, $i=1,2,3...,n$, $C_{Rij}$ is difference vector between geometric centroids of two point segments: $C_{Rij} = C(R_i) - C(R_j)$; $N_V\{R_i\}$ is the normal vector of point segments; $DistThresh$ is predefined distance threshold for determining adjacency in the graph that whether two point segments are connected by edge or not.

A second key step is to define the dissimilarity metric of adjacent region as the distance between two adjacent regions. To define this measure between neighboring regions, we first define an appropriate feature space. The point segments of laser data allow us to use several geometric, shape and physical properties to construct a joint spatial-physical feature space, which can adequately provide the most significant features for the segmentation of urban areas.

By defining the feature space, we can compute the weight matrix W of all point segments. The weight $w(u,v)$ between regions $u$ and $v$ is defined as

$$w(u,v) = \begin{cases} e^{-dI(u,v)*s}\cdot e^{-dZ(u,v)*s}\cdot e^{-dP(u,v)*s}\cdot e^{-dO(u,v)*s} & \text{if } u,v \text{ adjacent} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

with $dI(u,v) = \left|I(u)-I(v)\right|^2/\sigma_I$, $dZ(u,v) = \left|Z(u)-Z(v)\right|^2/\sigma_Z$, $dP(u,v) = \left|P(u)-P(v)\right|^2/\sigma_P$, $dO(u,v) = \left|O(u)-O(v)\right|^2/\sigma_O$

where $I(u)$, $Z(u)$, $P(u)$, $O(u)$ are mean intensity, vertical height of centroid, planarity, omnivariance [16] of each point segment, respectively. In addition, $s$ is a positive value controlling the sensitivity of each factor.

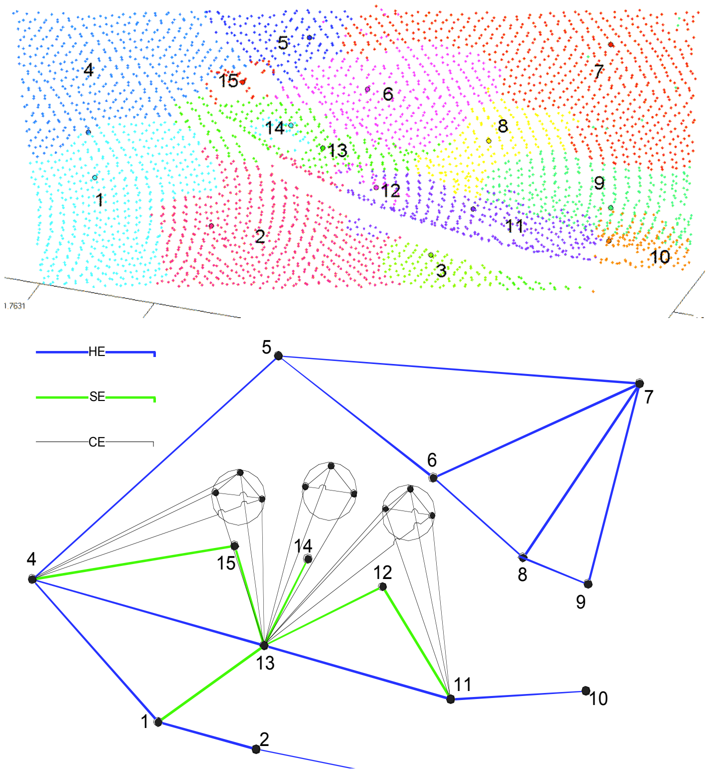



Figure.3. Top: Labeled point segments for a section (red dotted rectangle) of Fig.2a, bottom: non-planar RAG with multi child-nodes (within each black circular, for point segments 12, 14, 15, respectively) produced by spatial relation of nodes, with a point segment corresponding to a node.

Usually, Ncuts algorithms try to find a "balanced partition" [14][15] of a weighted graph via recursive cuts, which does not have a bias in favor of cutting small sets of isolated nodes in the graph. Unfortunately, some vehicles or trees in our case belong to this class, which are considered as one node of the weighted graph. The multi-child nodes strategy is suggested to the graph node which is connected with neighboring ones mainly by the spatial edge ***SE***, splitting each into multiple child nodes. There are edges (***CE***) connecting all the child nodes of two adjacent regions (Fig.3b). The weights between the child nodes within a region are all one, whereas the ones between two adjacent regions are all the same and equal to the weight between these two regions. This yields a new weight matrix $W_c = W \otimes E_c$, where $\otimes$ denotes the Kronecker product operator, and $E_c$ is the $c\times c$ matrix with all unit entries. In this experiment, we set the number of child nodes in each region to c = 3. In Fig.4, the segmentation result for Fig.2a after performing Ncuts grouping is depicted. Single segments can be used to describe significative and distinct parts of large-size objects (e.g. flyover or building), whereas many small-scale segments towards local details (e.g. vehicle in parking lots) are also preserved.

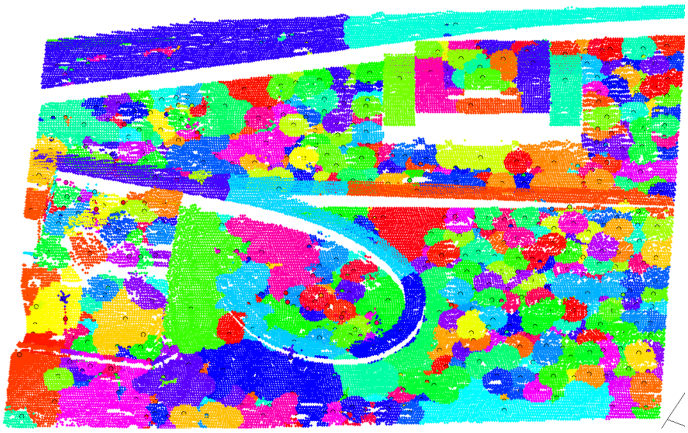


Figure 4. Segmentation result after using proposed Nucts grouping

## IV. CLASSIFICATION

Classification is performed on the set of point segments resulting from segmentation. The objective of this step is to verify the hypotheses for extracting two of urban object-classes: flyover and vehicle. A rule-based approach is used to do this task by examining various properties of point segments including following:

- ●elongation and width
- ●planarity and rectangularity
- ●vertical range

A hierarchical classifier following a course-to fine strategy is defined; where the attributes are employed to recursively partition the set of hypotheses into ever finer and more homogeneous subsets. The reason for doing this is that the urban area can be very heterogeneous, containing objects with a large variation in position, shape and reflectance signatures. It is easy for us to separate between large and small objects based on the length and width of the object at the highest level. We do this to separate large-sized man-made objects, e.g. roads and building parts, from average-sized objects, e.g. trees and vehicles. After the first partition, we can further perform parallel separations on two classes, respectively in order to distinguish different types of clutter objects from potential flyover and vehicles. Features that are used in this level include different geometric and topological features of point segments like width, elongation, rectangularity, planarity and vertical range. The test criteria at each level is set to avoid discarding potential interest objects

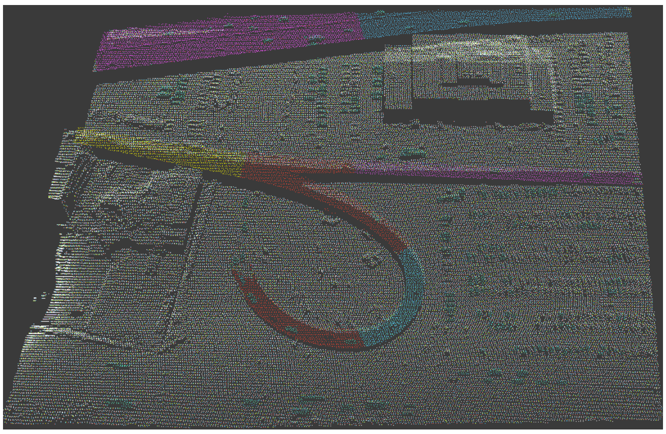
## V. EXPERIMENTS AND DISCUSSION

We have applied the proposed algorithm to two nadir ALS datasets over dense urban areas (Toronto city) for the extraction of urban objects. The test data feature complex structures with strong background clutters and a relatively low point density of ca.1.8 pts/$m^2$ [9]. Each comprises more than 300000 points and covers at least $350\times450m^2$ ground areas. As two examples, the experiments focused on the extraction of vehicle and elevated roads (flyovers or bridge in the city), which were rarely handled or regarded as inelegantly solved in former researches, moreover, they are objects at two totally different scales. It has been indicated that the bridges or flyovers in the urban areas cannot be considered as a part of normal roads that usually lie upon the terrain, and the vehicles are easily polluted by surrounding clutter points, e.g. vegetation, traffic sign or even road surface, etc [10]. Since one of the goals is to assess the semantic description capability of 3d point cloud, only the 3d geometric coordinates of ALS data were used to construct the feature space for MS with bandwidth $h=(hh)=(8, 6)$, and then physical and geometric attributes of resulted point segments are jointly considered in edge affinity matrix analyzed by the Ncuts.

Fig.5 shows us the final results. For verification of the classification results, it is necessary to determine the class of all the interest objects manually. For small-scale objects like vehicle, the appearance of the observed objects can be quite ambiguous in ALS data, even manual interpretation is difficult and its results cannot be referred to "ground-truth". The results of the automatic extraction were then assessed by comparing with the reference data. Two object-based accuracy measures were used: absolute/relative accuracy for object extraction AAOE and EAOE

$$\text{AAOE}=\frac{N_{1,1}}{N_r} \qquad \text{EAOE}=\frac{N_{1,1}}{N_s} \tag{6}$$

where $N_{1,1}$ is the number of objects which has one-to-one relationship with the reference data, $N_r$ is the true number of objects in the field, $N_s$ is the total number of automatically extracted objects. One-to-one relationship means that the overlaps between the boundaries of reference object and extracted object are greater than 80%. The quantitative results are summarized in Tab I, II. It can be inferred that flyover has been completely extracted and generally gained a better result than vehicle, whose results are mainly degraded by densely distributed vehicle instances (e.g. parking lots). Our algorithm is superior in extracting isolated vehicles with a high accuracy, while a model-based approach seems to be more competent for vehicles which are placed extremely close to each other [9].



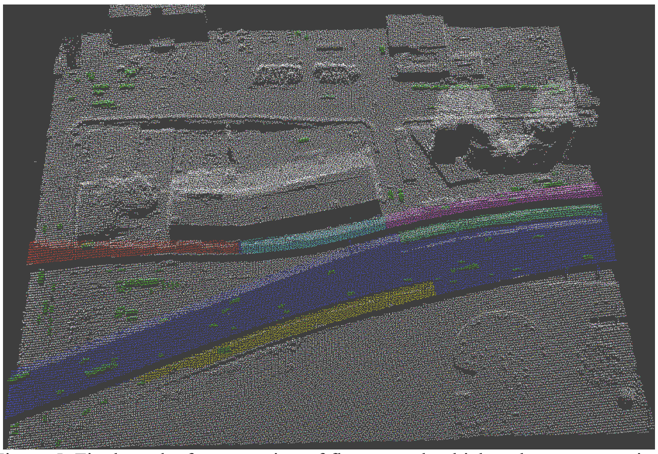


Figure.5. Final results for extraction of flyover and vehicle, where green points indicate vehicles and flyovers consisting of individual segments are marked by other different colors. Top: Toronto data 1, bottom: Toronto data 2

Note that MS is more feasible for parallel operations than the Ncuts method due to the complexity of eigendecomposition problem in the Ncuts. In the proposed method, the main computational cost is thus considerably reduced, since the number of point segments is by far smaller than that of the points. For first test dataset, the number of the point segments produced by MS is less than 600 regions (i.e., the RAG has less than 600 nodes). In addition, our approach can achieve an improved segmentation of ALS data towards semantic object analysis, because MS is able to generate more robust features considered for graph nodes compared to that used for the point nodes.

Tab I. EVALUATION FOR ELEVATED ROAD

| Dataset | AAOE | EAOE |
|---|---|---|
| Toronto 1 | 100% | 92.1% |
| Toronto 2 | 100% | 90.2% |

Tab II. EVALUATION FOR VEHICLE

| Dataset | AAOE | EAOE |
|---|---|---|
| Toronto 1 | 56.7% | 87.2% |
| Toronto 2 | 70.3% | 88.4% |

## VI. CONCLUSION

In this paper, a classification-based method is proposed to extract the flyover and vehicle in urban areas from ALS data based on 3d segmentation combining nonparametric clustering with spectral graph one. The approach started with genuine segmentation working directly on 3d raw laser data, retrieving inherently 3d structure primitives of urban areas. Then, the point segments were grouped into significative entities considering global optimization criterion based on a modified normalized cuts method. The semantic description capability of LiDAR data is assessed by the experiments on two datasets of inner-city to extract the flyover and vehicle. The results are very promising for flyover and isolated vehicles which were separated from strong clutters; only for densely distributed vehicles like in parking lots, the error rate is higher. Additionally, the integration of MS and Ncuts made the approach more practicable by reducing the computational cost greatly and more accuracy by incorporation more robust features derived from point segments. It can be claimed that both of MS and Ncuts methods work in a mutually enhanced fashion towards intelligent segmentation of complex ALS data.

## REFERENCES


[1] F. Rottensteiner, “Automatic generation of high-quality building models from LIDAR Data,” In: *IEEE Computer Graphics & Applications*, 23(6), pp. 42-50, 2003.

[2] J. Reitberger, P. Krzystek, U. Stilla, “Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees,” *International Journal of Remote Sensing*, 29(5): 1407-1431, 2008.S.

[3] P.J. Besl, R. C. Jain, “Segmentation through variable-order surface fitting,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10, 167–191, 1988.L. Koskeleinen.

[4] XY. Jiang, U. Meier, H. Bunke, “Fast Range Image Segmentation Using High-Level Segmentation Primitives,” In*: Proceedings of the 3rd IEEE Workshop on Applications of Computer Vision*, Washington, DC, P.83, 1996.

[5] R. Schnabel, R. Wahl and R. Klein, “Efficient RANSAC for Point-Cloud Shape Detection,” Computer Graphics Forum, 26(2), 214-226, 2007.R.M.

[6] G. Vosselman, B.G.H. Gorte, G. Sithole, T. Rabbani, “Recognizing structure in laser scanner point clouds,” *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 46, part 8/W2, Freiburg, Germany, October 4-6, pp. 33-38, 2004.

[7] S. Filin, N. Pfeifer, “Segmentation of airborne laser scanning data using a slope adaptive neighborhood,” *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(2), pp. 71-80, 2006.

[8] M. Rutzinger, B. Höfle, M. Hollaus and N. Pfeifer, “Object-Based point cloud analysis of full-waveform airborne laser scanning data for urban vegetation classification,”. *Sensors*, 8, 4505 – 4528, 2008.

[9] W. Yao, S. Hinz, U.Stilla, “Automatic vehicle extraction from airborne LiDAR data of urban areas using morphological reconstruction,” *5th IAPR Workshop on Pattern Recognition in Remote Sensing*, Florida, 2008.

[10] F. Clode, F, Rottensteiner, E, Kootsookos, Zelniker, “Detection and vectorisation of roads from LIDAR data,” In: *Photogrammetric Engineering & Remote Sensing*, 73(5), pp. 517-536, 2007.

[11] G. Sithole and G. Vosselman, “Bridge detection in airborne laser scanner data,” *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(1), Pages 33-46, 2006.

[12] T. Melzer, “Non-parametric segmentation of ALS point clouds using mean shift,” *Journal of Applied Geodesy*, 1(3): 159 – 170, 2007.

[13] D. Comaniciu, P. Meer, “Mean shift: a robust approach toward feature space analysis,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. Volume 24, Issue 5, Page(s):603 – 619, 2002.

[14] S. X. Yu and J. Shi, “Multiclass spectral clustering,” in Proc. Int. Conf.Comput. Vis., pp. 313–319, 2003.

[15] WB. Tao H. Hai and YM. Zhang, “Color Image Segmentation Based on Mean Shift and Normalized Cuts,” *Systems, Man, and Cybernetics, Part B, IEEE Transactions on*, 37(5), Pages 1382-1389, 2007.

[16] H. Gross, B. Jutzi, U. Thoennessen, “Segmentation of Tree regions Using data of a Full-Waveform Laser,” *Symposium of ISPRS Photogrammetric Image Analysis PIA07*, Vol XXXVI, Part 3/W49A. 2007