Segmentation and Classification of 3D Urban Point Clouds: Comparison and Combination of Two Approaches

A.K. Aijazi, A. Serna, B. Marcotegui, P. Checchin and L. Trassoudaine

Abstract Segmentation and classification of 3D urban point clouds is a complex task, making it very difficult for any single method to overcome all the diverse challenges offered. This sometimes requires the combination of several techniques to obtain the desired results for different applications. This work presents and compares two different approaches for segmenting and classifying 3D urban point clouds. In the first approach, detection, segmentation and classification of urban objects from 3D point clouds, converted into elevation images, are performed by using mathematical morphology. First, the ground is segmented and objects are detected as discontinuities on the ground. Then, connected objects are segmented using a watershed approach. Finally, objects are classified using SVM (Support Vector Machine) with geometrical and contextual features. The second method employs a super-voxel based approach in which the 3D urban point cloud is first segmented into voxels and then converted into super-voxels. These are then clustered together using an efficient link-chain method to form objects. These segmented objects are then classified using local descriptors and geometrical features into basic object classes. Evaluated on a common dataset (real data), both these methods are thoroughly compared on three different levels: detection, segmentation and classification. After analyses, simple strategies are also presented to combine the two methods, exploiting their complementary strengths and weaknesses, to improve the overall segmentation and classification results.

1 Introduction

The segmentation and classification of 3D point clouds for the interpretation of urban scenes and detailed semantic analysis have gained major interest in recent years. This considerable attention is due to the recent advancements in 3D data acquisition

Ahmad K. Aijazi, Paul Checchin and Laurent Trassoudaine
Université Blaise Pascal, Institut Pascal, BP 10448, F-63000 Clermont-Ferrand, France – CNRS, UMR 6602, Institut Pascal, F-63171 Aubière, France, e-mail: kamalaijazi@gmail.com, paul.checchin@univ-bpclermont.fr, laurent.trassoudaine@univ-bpclermont.fr
Andrés Serna and Beatriz Marcotegui
MINES ParisTech, CMM – Centre de Morphologie Mathématique, 35 rue St. Honoré, 77305 Fontainebleau-Cedex, France, e-mail: andres.serna_morales@mines-paristech.fr, beatriz.marcotegui@mines-paristech.fr

technologies as well as the increasing demand for different robotics applications in the field or service industry. Presenting a fundamental problem in robotics and computer vision, different research activities pertaining to automatic interpretation of 3D urban point clouds for various field robots and autonomous vehicles operating in outdoor environments are underway such as urban accessibility analysis [23], drivable road detection [4] and point cloud classification [17].

For scene interpretation and assignment of a semantic label to each 3D point (e.g. building, ground, trees, etc.), the first step is to segment the 3D point cloud. Point cloud segmentation can support classification and further feature extraction provided that the segments are logical groups of points belonging to the same object class. Some methods, including [20, 27], employ the use of small sets of specialized features, such as local point density or height from the ground, to discriminate only few object categories in outdoor scenes, or to separate foreground from background while some segmentation methods based on surface discontinuities, such as in [15], use surface convexity in a terrain mesh as a separator between objects. Lately, segmentation has been commonly formulated as graph clustering [9, 21]. Instances of such approaches are Graph-Cuts including Normalized-Cuts and Min-Cuts. Markov Random Fields are also used to segment and label 3D point clouds [2]. Different methods, such as in [17], are introduced in order to increase their efficiency while reducing their computational time.

The next step is to extract corresponding features from the segmented 3D object. These features rely on a local 3D neighborhood which is typically chosen as a spherical neighborhood formed by a fixed number of the k closest 3D points [13], spherical neighborhood with fixed radius [12] or cylindrical neighborhood with fixed radius [7]. These features are mainly based on geometrical features (shape, size, etc.) [19], local descriptors (color, intensity, surface normals, etc.) [1] or contextual features (position with respect to neighbors, etc.) [24].

Once these features have been calculated, the next step is the classification of each 3D point. Some methods such as [19, 1] rely on pre-defined geometrical models and thresholds but classification may also be conducted via different supervised learning techniques as well, such as Support Vector Machines [22], Gaussian Mixture Models [11], Random Forests [5], AdaBoost [14] and Bayesian Discriminant Classifiers [16]. Furthermore, contextual learning approaches also utilize relationships between 3D points in a local neighborhood which is usually inferred from the training data. Such methods for classifying point cloud data have been proposed with Associative and non-Associative Markov Networks [26], Conditional Random Fields [18] and multi-stage inference procedures focusing on point cloud statistics and relational information over different scales [28], etc. In addition to the above methods, Stamos *et al.* [8] propose an online algorithm to classify scanned points into 6 distinct classes (ground vegetation, car, horizontal surfaces, vertical surfaces and curb regions) during data acquisition by analyzing each scan-line one-by-one relying on several efficiently computed local features.

Common problems in this detection, segmentation and classification pipeline include coping with the complexity of 3D scenes caused by the irregular sampling, a large variety of objects, occlusions caused by obstructions, density variation caused

by different distances of objects from the sensors as well as the computational burden arising from large 3D point clouds and handling the various types of features. These diverse problems make it very difficult for any single method to produce the desired results. Hence, the combination of several approaches is necessary for different applications. Consequently, for effective combination, thorough evaluation and comparison is essential.

In this work, we present and compare two different approaches for segmenting and classifying 3D urban point clouds i.e. a method exploiting mathematical morphology (Sect. 2) and another based on super-voxels (Sect. 3). Evaluated on a common dataset (real data), both these methods are thoroughly compared (Sect. 4) on three different levels: detection, segmentation and classification. After analyses, simple strategies are also presented to combine the two methods, exploiting their complementary strengths and weaknesses, to improve the overall segmentation and classification results (Sect. 5).

2 Morphological Transformation Method

The method for segmenting 3D urban point cloud based on mathematical morphology is presented in [24]. It aims at developing a process to detect, segment and classify urban objects, suitable for large scale applications. In this method, the input point cloud is first mapped to a range image. This image is then interpolated in order to avoid connectivity problems and a k-flat zones algorithm is used to segment the ground (road and sidewalk). The facades and objects are extracted using morphological transformations. The method relies on facades being the highest vertical structures in the scene and objects are represented as bumps on the ground on the range image as shown in Fig. 1. Several geometrical and contextual features are computed for each object and classification is carried out using a standard SVM (Support Vector Machine). These features are summarized as follows:

- Geometrical features: object area and perimeter; bounding box area; maximum, mean, standard deviation and mode (the most frequent value) of the object height; object volume, computed as the integral of the elevation image over each object.
- Contextual features: Neighboring objects N_{neigh}, defined as the number of regions touching the object, using 8-connectivity on the elevation image. This feature is very discriminative in the case of group of trees and cars parked next to each other; confidence index $Cind = \frac{n_{real}}{n_{real}+n_{interp}}$, where n_{real} and n_{interp} are the number of non-empty object pixels before and after elevation image interpolation, respectively. In general, occluded and far objects have a low confidence index.

Relatively fast, the method uses little a priori information, and is based on robust morphological operators and supervised classification.

3 Super-voxel based Segmentation & Classification Method

This method presents a super-voxel based approach in which the 3D urban point cloud is first segmented into voxels and then converted into super-voxels. These

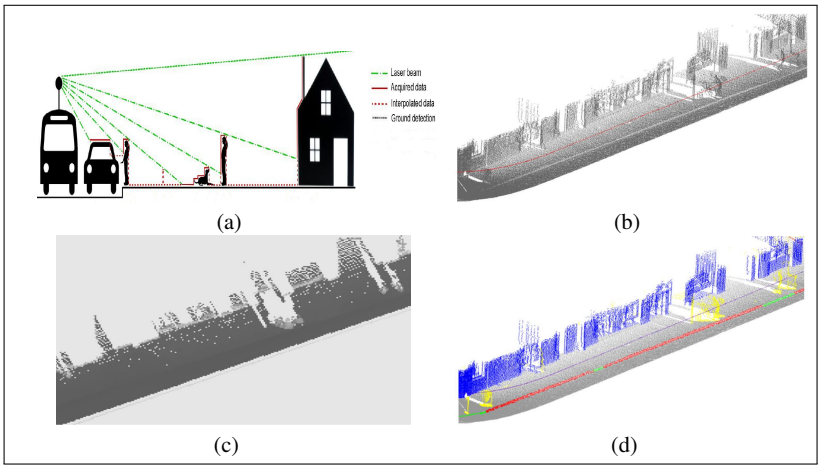


Fig. 1: (a) Segmentation of 3D point clouds based on morphological modeling. Objects are segmented out as bumps on the ground. (b) Input point cloud. (c) Range image. (d) Segmentation results.

are then clustered together using an efficient link-chain method to form objects. The method as presented in [1] uses an agglomerative clustering methodology to group 3D points based on r-NN (radius Nearest Neighbor). Although the maximum voxel size is predefined, the actual voxel sizes vary according to the maximum and minimum values of the neighboring points found along each axis to ensure the profile of the structure is maintained. A voxel is then transformed into a super-voxel when properties based on its constituting points are assigned to it. These properties mainly include: geometrical center, mean R, G&B value, maximum of the variance of R, G&B values; mean intensity value; variance of intensity values; voxel size along each axis X, Y&Z and surface normals of the constituting 3D points. With the assignment of all these properties, a voxel is transformed into a super-voxel. All these properties are then used to cluster these super-voxels into objects using a link chain method. In this method, each super voxel is considered as a link of a chain. All secondary links attached to each of these principal links are found. In the final step, all the principal links are linked together to form a continuous chain removing redundant secondary links in the process. If V_P be a principal link and V_n be the n^{th} secondary link then each V_n is linked to V_P if and only if the following three conditions are fulfilled:

$$\left|\mathbf{V}_{P_{X,Y,Z}} - \mathbf{V}_{n_{X,Y,Z}}\right| \leq (w_D + c_D)$$

$$\left|\mathbf{V}_{P_{R,G,B}} - \mathbf{V}_{n_{R,G,B}}\right| \leq 3\sqrt{w_C}$$

$$\left|\mathbf{V}_{P_I} - \mathbf{V}_{n_I}\right| \leq 3\sqrt{w_I}$$

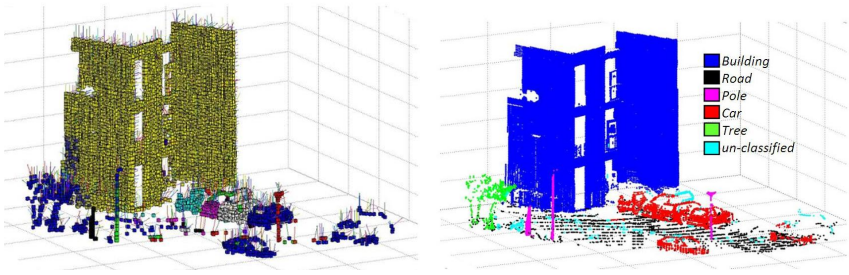
where, for the principal and secondary link super-voxels respectively:

- $\mathbf{V}_{P_{X,Y,Z}}$, $\mathbf{V}_{n_{X,Y,Z}}$ are the geometrical centers;

- $\mathbf{V}_{P_{R,G,B}}$, $\mathbf{V}_{n_{R,G,B}}$ are the mean R, G & B values;
- $\mathbf{V}_{P_I}$, $\mathbf{V}_{n_I}$ are the mean laser reflectance intensity values;
- w_C is the color weight equal to the maximum value of the two variances $Var(R,G,B)$, i.e. $\max(\mathbf{V}_{P_{Var(R,G,B)}}, \mathbf{V}_{n_{Var(R,G,B)}})$;
- w_I is the intensity weight equal to the maximum value of the two variances $Var(I)$.

w_D is the distance weight given as $\frac{\left(\mathbf{V}_{P_{s_{X,Y,Z}}} + \mathbf{V}_{n_{s_{X,Y,Z}}}\right)}{2}$. Here $s_{X,Y,Z}$ is the voxel size along X, Y & Z axis respectively. c_D is the inter-distance constant (along the three dimensions) added depending upon the density of points and also to overcome measurement errors, holes and occlusions, etc.

These clustered objects are then classified using local descriptors and geometrical features into 6 main classes: {Road, Building, Car, Pole, Tree, Unclassified}. These mainly include: surface normals, geometrical and barycenter, color, intensity, geometrical shape and size. The details of this method are presented in [1] while some results of this method are shown in Fig. 2. The salient features of this method are data reduction, efficiency and simplicity of approach.



(a) Voxelisation and segmentation into objects. (b) Labeled points.

Fig. 2: (a) Super-voxel based segmentation. (b) Classified 3D points.

4 Comparison: Results, Evaluation & Discussion

In order to compare the two approaches, we evaluated the two methods using the "Paris-Rue-Madame" dataset as presented in [25]. This database, used for benchmarking urban detection-segmentation-classification methods, consists of annotated 3D point clouds acquired by mobile terrestrial data acquisition system [10] of "Rue Madame" in the 6th Parisian district (France).

The evaluation was conducted for five common classes: {Building, Road, Pole, Tree, Car}. The detailed assessment carried out for each of the detection, segmentation and classification phase respectively are presented below.

4.1 Detection Evaluation

The detection evaluation is done to measure the capacity of the method to detect the objects present in the scene. This requires the choice of a criterion to decide if an object from the ground truth is detected or not. In order to ensure that this criterion does not bias the evaluation, the results are evaluated for a varying threshold m on the minimum object overlap as presented in [3]. In this analysis, an object OBJ is defined by the subset of points with the same object identifier i.e. S_{GT} and S_{AR} are the ground truth and the evaluated algorithm result subsets respectively. For any object j, S_{AR}^{j} is only validated as a correct detection of S_{GT}^{j} (a match) if the following condition is satisfied:

$$OBJ^{J}(detected) \xrightarrow{\text{iff}} \Big(\frac{|S_{GT}|}{|S_{GT} \cup S_{AR}|} > m\Big) \bigwedge \Big(\frac{|S_{AR}|}{|S_{GT} \cup S_{AR}|} > m\Big) \tag{1}$$

where $|.|$ is the cardinal (number of objects) of a set. The standard `Precision` `Pr` and `Recall` `Re` are then calculated as functions of m:

$$\texttt{Pr}(m) = \frac{\texttt{number\_of\_detected\_objects\_matched}}{\texttt{total\_number\_of\_detected\_objects}}$$
$$\texttt{Re}(m) = \frac{\texttt{number\_of\_detected\_objects\_matched}}{\texttt{total\_number\_of\_ground\_truth\_objects}}$$

These values of `Pr` and `Re` are then combined together to calculate the F-Measure as a function of m as expressed in Equation (2).

$$F(m) = 2 \times \frac{\texttt{Pr}(m) \times \texttt{Re}(m)}{\texttt{Pr}(m) + \texttt{Re}(m)} \tag{2}$$

Figure 3 shows the values of F-Measure with the variation of m for the different object types using both methods. The value of F-Measure decreases with the increasing value of m and this decay indicates the performance quality of the detection (good performance implies slower decay). Although the super-voxel based method does not classify motorcycles, they were detected and classified manually to analyze their segmentation quality (discussed in the next section).

The results show that the building and ground are much better detected by the morphological transformation method while the detection quality performance for cars, poles, and other road furniture is much more superior for the super-voxel based method.

4.2 Segmentation and Classification Evaluation

The evaluation was conducted for five common classes: {Building, Road, Pole, Tree, Car} and also the motorcycle class (only segmentation results). The segmentation and classification results are presented in Fig. 5. As trees were not present in the dataset, they were not considered for analysis.

The segmentation and classification quality was evaluated point-wise i.e. the number of 3D points correctly classified as members of a particular class. The re-

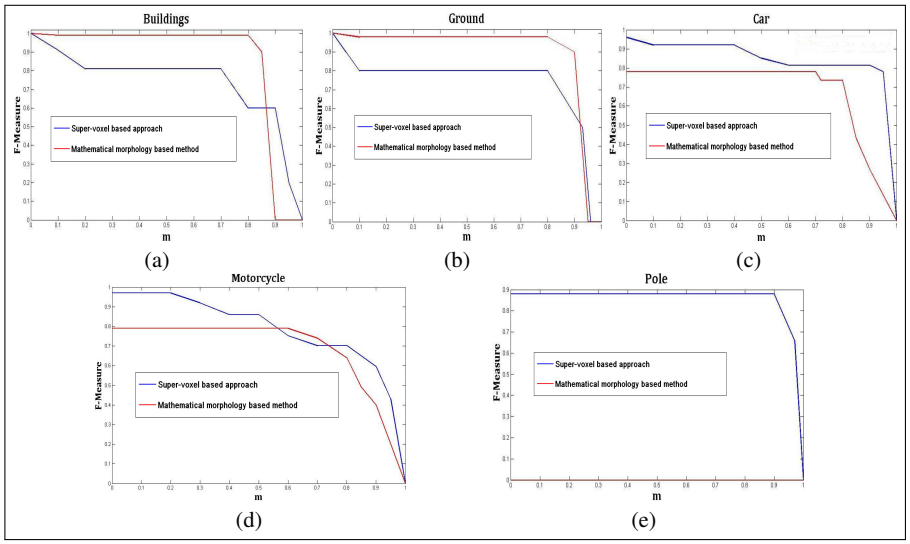


Fig. 3: Detection results for both super-voxel based and morphological transformation based methods are presented for 5 different classes in (a)-(e) respectively.

sults presented in Table 1 are in the form of a confusion matrix in which rows and columns are the class labels from the ground truth and the evaluated method respectively. The matrix values are the percentage of points with the corresponding labels using the metrics defined in [1]. If T_i, $i \in \{1,\cdots,N\}$, is the total number of 3D points distributed into objects belonging to N number of different classes in the ground truth and, and let t_{j_i}, $i \in \{1,\cdots,N\}$, be the total number of 3D points classified as a particular class of type-j and distributed into objects belonging to N different classes (for example a 3D point classified as part of the building class may actually belong to a tree) then the ratio S_{jk} (j is the class type as well as the row number of the matrix and $k \in \{1,\cdots,N\}$) is given as: $S_{jk} = \frac{t_{jk}}{T_k}$.

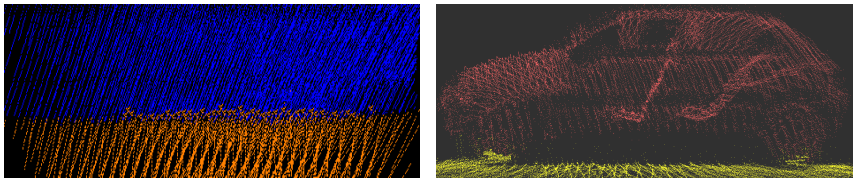
These values of S_{jk} are calculated for each class and are used to fill up each element of the confusion matrix, row by row. The Segmentation ACCuracy (**SACC**) is represented by the diagonal of the matrix while the values of classification accuracy (**CACC**), overall segmentation accuracy (**OSACC**) and overall classification accuracy (**OCACC**) are calculated as explained in [1].

Compared to contemporary evaluation methods such as used in [17], employing a standard confusion matrix, this method is more suitable for this type of work as it provides more insight in segmentation results along with the classification results and directly gives the segmentation accuracy similar to [6]. Also as compared to standard precision and recall evaluation, the use of this metric, also accommodates for the unclassified 3D points in the results giving a more accurate result without incorporating the unclassified objects as a class in the confusion matrix.

Table 1 shows the results. It can be seen that for the super-voxel based method, some of the 3D points belonging to different object classes are found in the road

Table 1: Segmentation and classification results for both super-voxel based and mathematical morphology-based method (inside braces) are presented respectively.

	Building	Road	Pole	Car	**CACC**
Building	0.914 (0.986)	0.013 (0.045)	0.000 (0.000)	0.000 (0.010)	0.950 (0.970)
Road	0.02 (0.002)	0.901 (0.940)	0.005 (0.000)	0.010 (0.002)	0.933 (0.968)
Pole	0.000 (0.000)	0.001 (0.001)	0.710 (0.000)	0.000 (0.010)	0.850 (0.495)
Car	0.000 (0.010)	0.005 (0.195)	0.000 (0.000)	0.900 (0.950)	0.950 (0.870)
Overall segmentation accuracy: **OSACC**				0.856 (0.720)	
Overall classification accuracy: **OCACC**					0.920 (0.825)



(a) (b)

Fig. 4: (a) and (b) show the misclassification of some 3D points at boundary regions of road surface with building and car respectively for super-voxel based method.

class and vice versa. This was found evident at boundary regions of objects belonging to two different classes, as shown in Fig. 4, as sometimes in the voxelisation process, some of the 3D points belonging to adjacent objects are incorporated in the same voxel if they have similar color and reflectance intensity values.

Also, it was found that, for this method, one of the traffic sign post was wrongly classified as a tree resulting in a low **SACC** and **CACC** of 0.71 and 0.85 respectively. This was due to the fact that the particular sign post contained two traffic signs on the same post giving it a small tree like appearance (in 3D point cloud at least) as shown in Fig. 6. Compared to this method, the morphological transformation method failed to classify any of the poles correctly (as depicted in the table), confusing most of them with trees.

Also evident from the table, the interaction between classes is much more significant in the case of the morphological transformation method while on the other hand in the super-voxel based method the segmented objects belonging to a particular class instead of being distributed in other classes rather remain unclassified.

In order to further assess the quality of segmentation, the ratio (**f**) of the total number of objects segmented by the applied method and the total number of segmented objects in the ground truth was plotted for each of the object classes as shown in Fig. 7. A value of 1 represents overall best segmentation whereas a value greater than 1 denotes overall over-segmentation while a value less than 1 denotes overall under-segmentation. A value of 0 shows failure to detect or no detection.

The mathematical morphology based method seems to outperform the super-voxel based method in terms of segmenting building and road surface. In the super-voxel based segmentation method the road was over-segmented in 4 parts as they were found disconnected and also one of the building was over-segmented due to

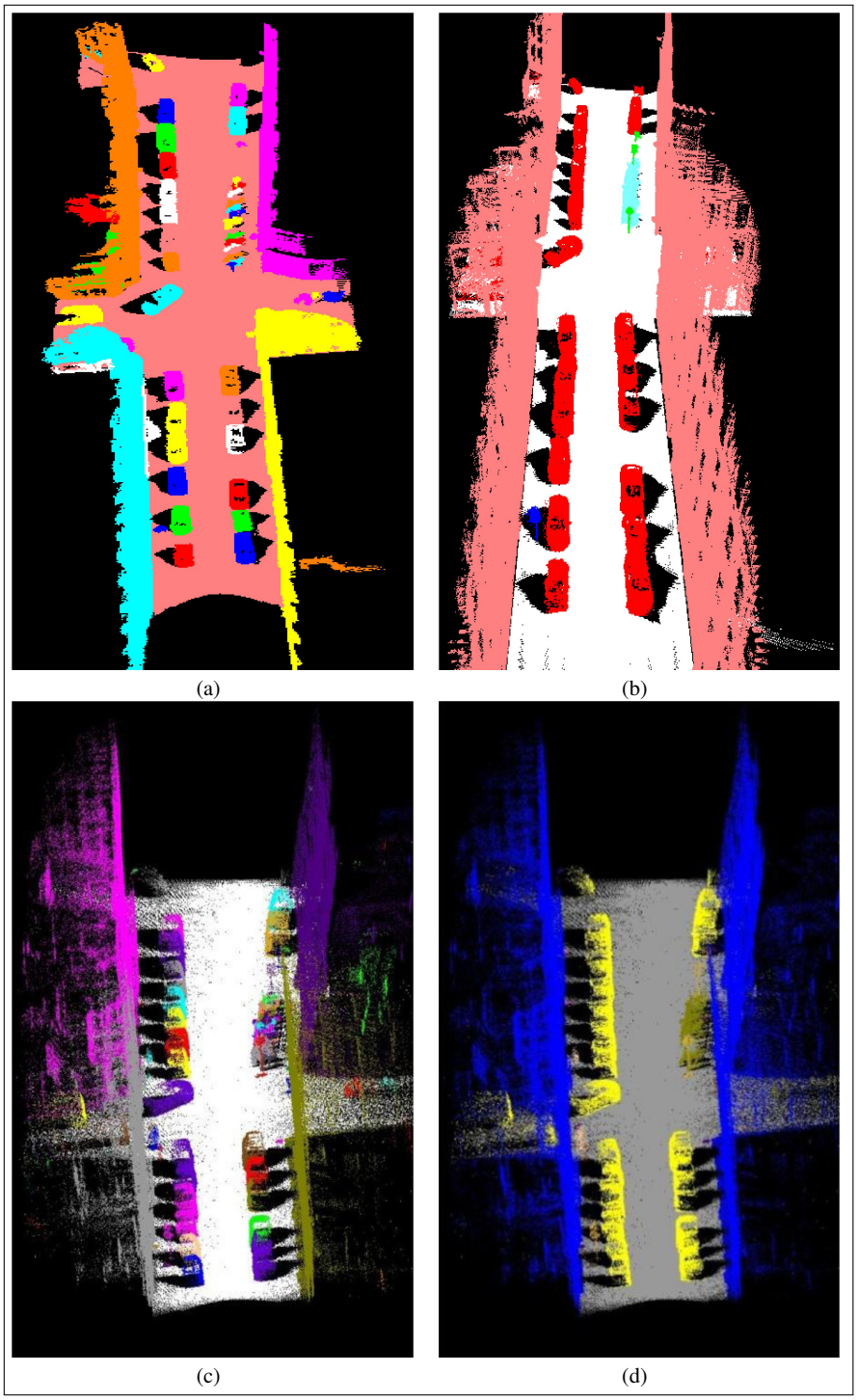
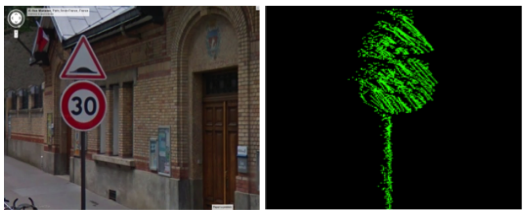


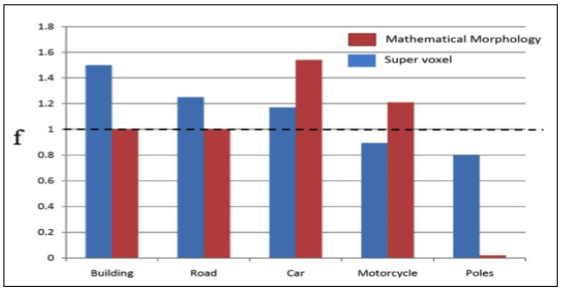
Fig. 5: (a) & (b) show the segmentation and classification results for super-voxel based method while (c) & (d) show the segmentation and classification results for morphological transformation based method respectively. In (a) & (c) every segmented object is represented by a separate color (some colors are repeated) while in (b) & (d) each class is represented by a different color.



(a) (b)

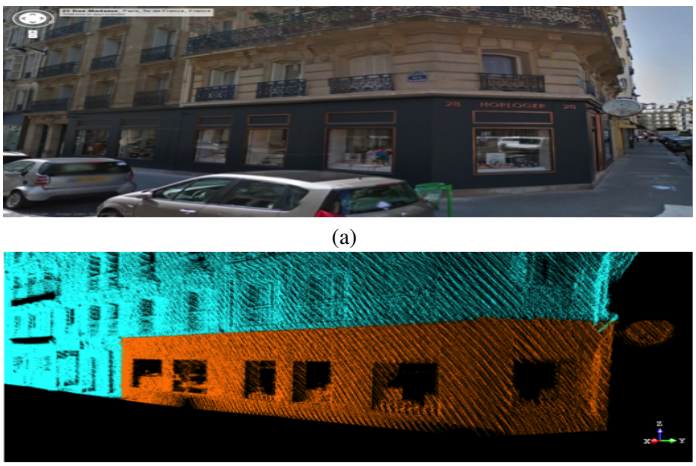
Fig. 6 (a) Google street view photo of the sign post with two traffic signs on Rue Madame. (b) Corresponding 3D points.

Fig. 7 Overall segmentation quality of the two methods for different object classes.



strong variation in color and reflectance intensity values (as shown in Fig. 8) while in case of another building small part found disjoint from the main building was segmented as a separate object (shown in Fig. 9).

However, compared to the mathematical morphology based method, the super-voxel based method segments cars and other road furniture better as apart from the adjacency of the 3D points it also uses color and reflectance intensity values in the segmentation phase. Figure 10 shows the segmentation results, for both methods,



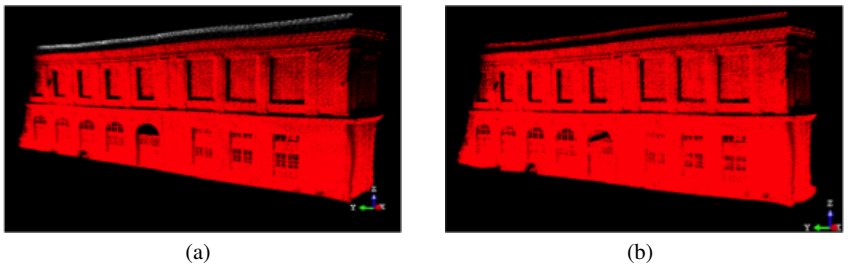
(a)

(b)

Fig. 8: (a) Google street map view photo of the building on Rue Madame with a strong variation of paint color. (b) Segmentation results of the super-voxel based method.

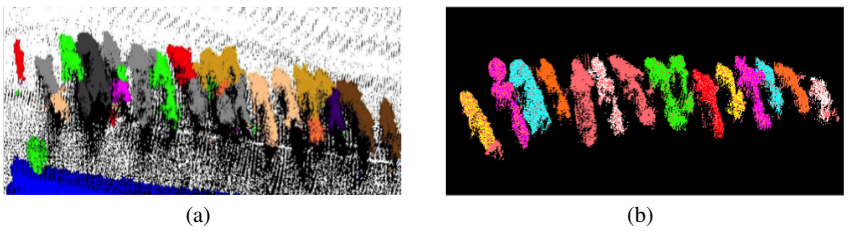
of some of the motorcycles parked in the scene. For the super-voxel based method, we also find in one instance that two cars parked very close together, having similar color and reflectance intensity values, are segmented out as one single car.

The mathematical morphology based method, constrained by the generated profile, also fails to segment out 3D ground points directly under the motorcycles and car as shown in Fig. 11. These ground points are hence considered as part of the car (also expressed in Table 1 i.e. value of 0.195). This is not an issue for the super-voxel based method relying on local descriptors i.e. color, reflectance intensity and surface normals.



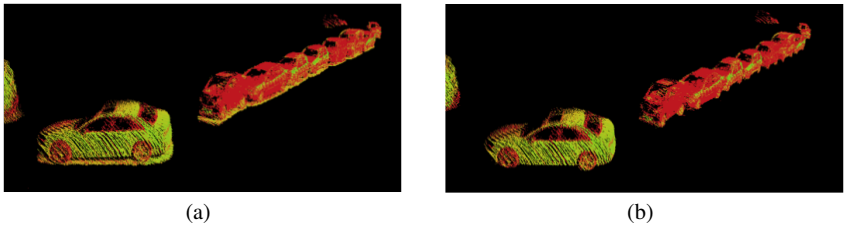
(a) (b)

Fig. 9: (a) and (b) show the segmentation results of a particular building in Rue Madame for super-voxel based method and mathematical morphology based method respectively. In (a) it can be seen that part of the building that was disjoint was segmented as a separate object.



(a) (b)

Fig. 10: (a) and (b) show the segmentation results of some motorcycles parked in the street for mathematical morphology and super-voxel based method respectively.



(a) (b)

Fig. 11: (a) and (b) show the segmentation results of some cars in the street for both mathematical morphology and super-voxel based methods respectively. (a) shows some ground point directly underneath the cars, segmented as part of the cars.

5 Combining The Two Approaches

In order to exploit the strengths of the two methods and overcome their respective weaknesses, we combined the results of the two approaches. Two different types of combinations were tried which are explained below.

5.1 Direct combination

In this combination, a simple union is applied to the segments, from the two methods, belonging to the same objects from the different object classes. A simple overlap ratio of 75% was set (i.e. if more than 75% of overlap between two segments, they are merged together as one). The improved results are presented in Table 2. We find that although the segmentation and classification results improve slightly (**OSACC** $= \mathbf{0.896}$, **OCACC** $= \mathbf{0.922}$), the overall segmentation quality decreases, due to the fact that combining of segments for each object class, in such a manner, often results in over-segmentation as shown in Fig. 12.

Table 2: Segmentation and classification results for direct combination are presented.

	Building	Road	Pole	Car	**CACC**
Building	0.986	0.031	0.000	0.010	0.972
Road	0.015	0.940	0.005	0.010	0.955
Pole	0.000	0.001	0.710	0.006	0.851
Car	0.001	0.110	0.000	0.950	0.912
Overall segmentation accuracy: **OSACC**				0.896	
Overall classification accuracy: **OCACC**					0.922

5.2 Selective combination

In order to preserve the strengths of each method and overcome their respective weaknesses, a selective combination is proposed. Using the complimentary performances of the two approaches as discussed in Sect. 1, we combine the outputs of the two methods i.e. mathematical morphology based method for building and road surface while super-voxel based method for other classes and road furniture. The improved results are presented in Table 3. We find that not only the segmentation and classification results improve (**OSACC** $= \mathbf{0.884}$, **OSACC** $= \mathbf{0.935}$), but also the segmentation quality as shown in Fig. 12.

6 Conclusion

In this paper, we present and compare two different approaches for segmenting and classifying of 3D urban point clouds i.e. one based on mathematical morphology while the other on super-voxels. Evaluated on a common dataset (real data), both these methods are thoroughly compared on three different levels: detection, segmentation and classification. The results show that the building and ground are much better detected by the mathematical morphology based method while the detection quality performance for cars, poles, and other road furniture is much more superior for the super-voxel based method. After analyses, simple strategies are also presented to combine the two methods, exploiting their complementary strengths and weaknesses, to improve the overall segmentation and classification results.

Table 3: Segmentation and classification results for selective combination are presented.

	Building	Road	Pole	Car	**CACC**
Building	0.986	0.045	0.000	0.000	0.970
Road	0.002	0.940	0.000	0.002	0.968
Pole	0.000	0.001	0.710	0.000	0.854
Car	0.000	0.002	0.000	0.900	0.950
Overall segmentation accuracy: **OSACC**				0.884	
Overall classification accuracy: **OCACC**					0.935

The same comparison methodology can be easily adapted to compare other segmentation and classification methods while the combination strategies need to be further studied and better adapted to improve upon the overall performances, for different applications.

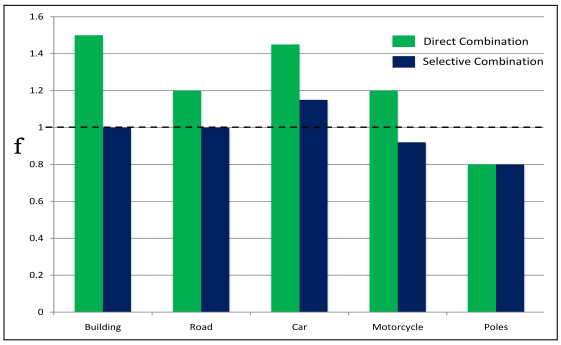


Fig. 12 Overall segmentation quality for different object classes, for the two combination methods.

Acknowledgements The work reported in this paper was supported by the French national research agency (ANR CONTINT iSpace & Time – ANR- 10-CONT-23) and has been performed as part of Cap Digital Business Cluster TerraMobilita project.

References

1. Aijazi, A.K., Checchin, P., Trassoudaine, L.: Segmentation Based Classification of 3D Urban Point Clouds: A Super-Voxel Based Approach. Remote Sensing **5**(4), 1624–1650 (2013)
2. Anguelov, D., Taskar, B., Chatalbashev, V., Koller, D., Gupta, D., Heitz, G., Ng, A.: Discriminative Learning of Markov Random Fields for Segmentation of 3D Scan Data. In: CVPR, IEEE Conf. on, vol. 2, pp. 169–176. Los Alamitos, CA, USA (2005)
3. Brédif, M., Vallet, B., Serna, A., Marcotegui, B., Paparoditis, N.: Terramobilita/IQmulus Urban Point Cloud Analysis Benchmark. In: IQmulus workshop in conjunction with SGP 14. Cardiff, UK (2014)
4. Byun, J., Na, K.i., Seo, B.s., Roh, M.: Drivable Road Detection with 3D Point Clouds Based on the MRF for Intelligent Vehicle. In: L. Mejias, P. Corke, J. Roberts (eds.) Field and Service Robotics, *Springer Tracts in Advanced Robotics*, vol. 105, pp. 49–60. Springer International Publishing (2015)
5. Chehata, N., Guo, L., Mallet, C.: Airborne lidar feature selection for urban classification using random forests. The Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **38**(3), 207–212 (2009)
6. Douillard, B., Underwood, J., Kuntz, N., Vlaskine, V., Quadros, A., Morton, P., Frenkel, A.: On the Segmentation of 3D LIDAR Point Clouds. In: IEEE Int. Conf. on Robotics and Automation (ICRA), p. 8. Shanghai, China (2011)

7. Filin, S., Pfeifer, N.: Segmentation of airborne laser scanning data using a slope adaptive neighborhood. {ISPRS} Journal of Photogrammetry and Remote Sensing **60**(2), 71–80 (2006)
8. Friedman, S., Stamos, I.: Online detection of repeated structures in point clouds of urban scenes for compression and registration. Int. J. of Computer Vision **102**(1-3), 112–128 (2013)
9. Golovinskiy, A., Funkhouser, T.: Min-Cut Based Segmentation of Point Clouds. In: IEEE Workshop on Search in 3D and Video (S3DV) at ICCV, pp. 39 – 46 (2009)
10. Goulette, F., Nashashibi, F., Abuhadrous, I., Ammoun, S., Laurgeau, C.: An integrated on-board laser range sensing system for on-the-way city and road modelling. ISPRS RFPT (2006)
11. Lalonde, J.F., Unnikrishnan, R., Vandapel, N., Hebert, M.: Scale selection for classification of point-sampled 3D surfaces. In: 5th Int. Conf. on 3-D Digital Imaging and Modeling, pp. 285–292 (2005)
12. Lee, I., Schenk, T.: Perceptual organization of 3d surface points. In: The Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXIV, Part 3A, pp. 193–198 (2002)
13. Linsen, L., Prautzsch, H.: Global versus local triangulations. In: J. Roberts (ed.) Proc. of Eurographics 2001, Short Presentations, pp. 257–263. Oxford, UK (2001)
14. Lodha, S., Fitzpatrick, D., Helmbold, D.: Aerial lidar data classification using adaboost. In: 3-D Digital Imaging and Modeling, 2007. 3DIM'07. 6th Int. Conf. on, pp. 435–442 (2007)
15. Moosmann, F., Pink, O., Stiller, C.: Segmentation of 3D lidar data in non-flat urban environments using a local convexity criterion. In: IEEE Intelligent Vehicles Symposium (IV), pp. 215–220 (2009)
16. Munoz, D., Bagnell, J.A.D., Vandapel, N., Hebert, M.: Contextual Classification with Functional Max-Margin Markov Networks. In: CVPR, IEEE Conf. on, pp. 975 – 982 (2009)
17. Munoz, D., Vandapel, N., Hebert, M.: Onboard contextual classification of 3-D point clouds with learned high-order Markov Random Fields. In: IEEE Int. Conf. on Robotics and Automation, pp. 2009–2016 (2009)
18. Niemeyer, J., Rottensteiner, F., Soergel, U.: Conditional Random Fields For Lidar Point Cloud Classification in Complex Urban Areas. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences **I-3**, 263–268 (2012)
19. Oude Elberink, S., Kemboi, B.: User-assisted object detection by segment based similarity measures in mobile laser scanner data. ISPRS - Int. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XL-3**, 239–246 (2014)
20. Rabbani, T., van den Heuvel, F.A., Vosselmann, G.: Segmentation of point clouds using smoothness constraint. In: IEVM06 (2006)
21. Schoenberg, J., Nathan, A., Campbell, M.: Segmentation of dense range information in complex urban scenes. In: IEEE/RSJ Int. Conf. on Intellig. Robots and Systems (IROS), pp. 2033 – 2038. Taipei, Taiwan (2010)
22. Secord, J., Zakhor, A.: Tree detection in urban regions using aerial lidar and image data. Geoscience and Remote Sensing Letters, IEEE **4**(2), 196–200 (2007)
23. Serna, A., Marcotegui, B.: Urban accessibility diagnosis from mobile laser scanning data. {ISPRS} Journal of Photogrammetry and Remote Sensing **84**(0), 23 – 32 (2013)
24. Serna, A., Marcotegui, B.: Detection, segmentation and classification of 3d urban objects using mathematical morphology and supervised learning. {ISPRS} Journal of Photogrammetry and Remote Sensing **93**(0), 243 – 255 (2014)
25. Serna, A., Marcotegui, B., Goulette, F., Deschaud, J.E., et al.: Paris-Rue-Madame database: a 3D mobile laser scanner dataset for benchmarking urban detection, segmentation and classification methods. In: 4th Int. Conf. on Pattern Recognition, Applications and Methods (2014)
26. Shapovalov, R., Velizhev, A., Barinova, O.: Non-associative markov networks for 3d point cloud classification. In: Photogrammetric Computer Vision and Image Analysis (PCV 2010), vol. 38, pp. 103–108 (2010)
27. Sithole, G., Vosselman, G.: Automatic structure detection in a point-cloud of an urban landscape. In: 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, pp. 67–71 (2003)
28. Xiong, X., Munoz, D., Bagnell, J.A.D., Hebert, M.: 3-D Scene Analysis via Sequenced Predictions over Points and Regions. In: IEEE Int. Conf. on Robotics and Automation (2011)