HAL
open science

Paris-Lille-3D: A Point Cloud Dataset for Urban Scene Segmentation and Classification

Xavier Roynard, Jean-Emmanuel Deschaud, François Goulette

To cite this version:

Xavier Roynard, Jean-Emmanuel Deschaud, François Goulette. Paris-Lille-3D: A Point Cloud Dataset for Urban Scene Segmentation and Classification. CVPR Workshop on Real-World Challenges and New Benchmarks for Deep Learning in Robotic Vision, Jun 2018, Salt Lake City, United States. hal-01959556

HAL Id: hal-01959556

https://hal.science/hal-01959556

Submitted on 18 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Paris-Lille-3D: A Point Cloud Dataset for Urban Scene Segmentation and Classification

Xavier Roynard
xavier.roynard@mines-paristech.fr

Jean-Emmanuel Deschaud
jean-emmanuel.deschaud@mines-paristech.fr

François Goulette
francois.goulette@mines-paristech.fr

MINES ParisTech, PSL Research University, Centre for Robotics

Abstract

This article presents a dataset called Paris-Lille-3D[1]*. This dataset is composed of several point clouds of outdoor scenes in Paris and Lille, France, with a total of more than 140 million hand labeled and classified points with more than 50 classes (e.g., the ground, cars and benches). This dataset is large enough and of high enough quality to further research on techniques regarding the automatic classification of urban point clouds. The fields to which that research may be applied are vast, as it provides the ability to increase productivity in regards to the management of urban infrastructures. Moreover, this type of data has the potential to be crucial in the field of autonomous vehicles.*

1. Introduction

LiDAR mobile mapping systems were developed in the early 2000s ([3], [1]). Since then, these systems also called Mobile Laser Scanning (MLS) have improved in regards to robustness. They are now marketed by various companies (such as Riegl and Optech) and are increasingly used by surveyor companies. MLS, unlike fixed LiDAR significantly improves productivity and makes it possible to digitize entire cities in 3D. One example is the dataset TorontoCity [10] with LiDAR data on the entire city of Toronto (with high-precision maps as ground truth but without ground truth regarding point wise classification).

Key aspect of MLS technology is the ability to process large volumes of product data, including segmentation and classification steps. Segmentation is the ability to group points into homogeneous regions. Classification is the ability to give a semantic class to a group of points or to each point (e.g., car, bench and tree).

The two main fields of LiDAR application that require 3D data semantics are infrastructure's management by cities and the creation of precise maps for autonomous vehicles. Cities are increasingly using MLS scans to inventory their street furniture and to update their Geographic Information Systems (GIS). MLS's production capacity allows cities to update their GISs at higher frequencies and with greater accuracies (measurement standardizations) than other systems.

Another major application of 3D outdoor scene classification is the ability to create maps for the navigation of autonomous vehicles. Segmented and classified point clouds can be used to locate vehicles ([5]) and to provide critical information regarding urban scenes (e.g., traffic lights, poles and traffic signs) to assit such vehicles with real-time navigation. Currently, work is underway on the ability of the infrastucture to send up-to-date point clouds to vehicles ([2]).

The Paris-Lille-3D dataset can further research on this topic and allow deep learning methods to be developed that involve large volumes of data.

2. Dataset

The Paris-Lille-3D dataset were acquired with the MLS prototype of the center for robotics of MINES ParisTech. The system is a van equipped with a precise localisation system (GPS Novatel FlexPak 6 + IMU Ixsea LANDINS) and a Velodyne HDL-32e LiDAR mounted at the rear of the truck. We have taken care to keep the origin of the Lidar for each measured point and its timestamp. The dataset comprises three point clouds files: Lille1.ply, Lille2.ply, Paris.ply. Each serve as a training base for tests on the two others. There are more than 140 million manually classified points and 50 classes. Indeed, the separation of classes was furthered to note vehicles parked, stopped on the road, and

[1] http://caor-mines-paristech.fr/fr/paris-lille-3d-dataset/

in motion. Figure 1 shows the different classes that were used in the data. The added difficulty is that some classes contain very few examples which makes the task of learning very challenging. The article [6] can be referred to for details on the different files and annotations that were used.

	Class	Number of samples (of points)							
		Lille1		Lille2		Paris		Total	
	unclassified	1	(54.88k)	1	(16.47k)	1	(60.79k)	3	(132.1k)
	other	16	(70.15k)	11	(14.10k)	11	(30.82k)	38	(115.1k)
surface	road	1	(34.80M)	1	(11.82M)	1	(19.68M)	3	(66.30M)
surface	sidewalk	18	(8.566M)	8	(4.97M)	5	(4.6M)	31	(16.67M)
surface	island	7	(458.8k)	0	(0)	9	(82.46k)	16	(541.2k)
surface	vegetation	32	(562.2k)	22	(512.4k)	6	(157.1k)	60	(1.232M)
surface	building	34	(18.2M)	8	(7.934M)	5	(9.431M)	47	(35.39M)
static	post	9	(13.51k)	0	(0)	0	(0)	9	(13.51k)
static	bollard	122	(34.80k)	64	(7.982k)	84	(28.33k)	270	(71.10k)
static	floor lamp	72	(232.9k)	13	(23.69k)	21	(113.2k)	106	(369.7k)
static	traffic light	14	(25.82k)	11	(27.41k)	16	(80.76k)	41	(133.10k)
static	traffic sign	82	(113.6k)	39	(69.89k)	17	(30.75k)	138	(214.3k)
static	signboard	20	(68.34k)	12	(43.14k)	3	(13.41k)	35	(124.9k)
static	mailbox	0	(0)	0	(0)	1	(4.739k)	1	(4.739k)
static	trash can	11	(14.72k)	6	(17.43k)	22	(30.35k)	39	(62.50k)
static	meter	0	(0)	0	(0)	2	(2.840k)	2	(2.840k)
static	bicycle terminal	10	(1.736k)	0	(0)	13	(6.451k)	23	(8.187k)
static	bicycle rack	8	(774)	0	(0)	14	(8.665k)	22	(9.439k)
static	statue	2	(4.158k)	0	(0)	0	(0)	2	(4.158k)
static	barrier	45	(83.94k)	5	(9.286k)	7	(56.10k)	57	(149.3k)
static	roasting	6	(831.0k)	1	(20.82k)	4	(1.677M)	11	(2.529M)
static	wire	34	(14.18k)	8	(9.325k)	0	(0)	42	(23.51k)
static	low wall	58	(840.2k)	2	(26.99k)	9	(951.4k)	69	(1.819M)
static	shelter	0	(0)	0	(0)	3	(83.45k)	3	(83.45k)
static	bench	5	(2.911k)	1	(364)	2	(1.391k)	8	(4.666k)
static	distribution box	19	(50.28k)	8	(14.89k)	3	(53.4k)	30	(118.2k)
static	lighting console	78	(7.251k)	60	(9.731k)	9	(4.393k)	147	(21.38k)
static	windmill	1	(10.17k)	0	(0)	0	(0)	1	(10.17k)
dynamic	pedestrian	17	(24.81k)	7	(11.60k)	61	(150.2k)	85	(186.6k)
dynamic	parked bicycle	15	(9.74k)	0	(0)	33	(81.67k)	48	(90.75k)
dynamic	mobile scooter	0	(0)	0	(0)	1	(131)	1	(131)
dynamic	parked scooter	0	(0)	0	(0)	31	(169.1k)	31	(169.1k)
dynamic	mobile motorbike	0	(0)	0	(0)	1	(1.613k)	1	(1.613k)
dynamic	parked motorbike	2	(2.428k)	0	(0)	4	(14.37k)	6	(16.79k)
dynamic	mobile car	21	(175.0k)	4	(40.96k)	5	(66.35k)	30	(282.3k)
dynamic	stopped car	0	(0)	1	(28.27k)	1	(9.375k)	2	(37.64k)
dynamic	parked car	182	(2.266M)	47	(853.7k)	65	(1.610M)	294	(4.730M)
dynamic	mobile van	3	(97.27k)	1	(41.6k)	0	(0)	4	(138.3k)
dynamic	parked van	5	(84.75k)	5	(85.20k)	9	(357.6k)	19	(527.6k)
dynamic	stopped truck	0	(0)	0	(0)	1	(235.7k)	1	(235.7k)
dynamic	parked truck	2	(40.32k)	0	(0)	1	(53.44k)	3	(93.76k)
dynamic	stopped bus	0	(0)	0	(0)	1	(78.41k)	1	(78.41k)
dynamic	parked bus	1	(9.623k)	0	(0)	0	(0)	1	(9.623k)
dynamic	table	0	(0)	0	(0)	2	(576)	2	(576)
dynamic	chair	0	(0)	0	(0)	8	(4.842k)	8	(4.842k)
dynamic	trash can	138	(148.8k)	80	(115.9k)	0	(0)	218	(264.7k)
dynamic	waste	5	(2.307k)	0	(0)	0	(0)	5	(2.307k)
natural	natural	149	(1.233M)	47	(396.9k)	36	(1.610M)	232	(3.240M)
natural	tree	72	(2.310M)	23	(565.10k)	101	(4.755M)	196	(7.631M)
natural	potted plant	32	(72.80k)	5	(26.96k)	0	(0)	37	(99.76k)
	Total	1349	(71.36M)	501	(26.84M)	629	(45.80M)	2479	(143.10M)

Figure 1. The classes used in the dataset.

For the dataset, great care was taken to ensure quality segmentations and classifications were used. Most other urban datasets (such as [8] and [9]) used semi-automatic methods to create ground truth, and this led to inaccuracies in regards to semantization that could block learning, especially for deep methods.

Figure 2 shows part of one of the point clouds in the dataset. It is raw data that was obtained using the mobile mapping system. The colors correspond to the intensity of the LiDAR return signal. Figure 3 shows a point cloud with labels corresponding to a manual segmentation of this same part of the point cloud. The numerous objects present in the urban scene are shown. Finally, Figure 4 shows the same point cloud with manual classification that can be used as ground truth.

The dataset closest to the Paris-Lille-3D dataset is the Semantic3D dataset [4]. It was formed using a set of acquisitions of fixed scanners and was also hand labeled. A benchmark with test data is available online[2]. However, the number of classes available in the dataset is relatively limited as it has only eight classes, unlike the Paris-Lille-3D dataset. Moreover, the data from the Semantic3D dataset is quite different than the data that can be gathered using mobile mapping systems. MLS systems are better suited to digitizing large areas, such as cities, from a terrestrial point of view, and thus, they should be popularized. More details about shortcomings of other existing benchmarks are presented in [6].

We propose two evaluation metrics : one for the object detection task and one for the point wise classification. To evaluate detection of objects, we use the same metric as used in iQmulus/TerraMobilita contest [9]. For the point wise classification, we take classical precision, recall, F1 score and IoU to evaluate performance [6].

The quality of the manual classification and the quantity of data in the Paris-Lille-3D dataset will enable deep-network learning. The publication [7] presents a new deep-network trained on the Paris-Lille-3D dataset. In addition to detecting numerous objects present in an urban scene, the quality of a good classification method lies in its ability to demonstrate that what works in one part of a city works not only in another part of the same city but also in a different city. This is why two different cities were used in the dataset: Paris and Lille in France. A classification benchmark for the dataset will soon be released based on non-published data in Paris, Lille and other cities.

3. Conclusion

We have presented in this article the Paris-Lille-3D dataset and the interest of having this type of dataset for the scientific community to improve the techniques of automatic classification of urban scenes. Important characteristics of our benchmark are high quality of segmentation and classification (the annotation of each point was done completely manually), the wide variety of object classes and that the dataset is quite large, making it useful for deep learning techniques.

References

[1] I. Abuhadrous, S. Ammoun, F. Nashashibi, F. Goulette, and C. Laurgeau. Digitizing and 3d modeling of urban environments and roads using vehicle-borne laser scanner system. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, volume 1, pages 76–81 vol.1, Sept 2004.

[2] K. F. Chu, E. R. Magsino, I. W. H. Ho, and C. K. Chau. Index coding of point cloud-based road map data for autonomous driving. In *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, pages 1–7, June 2017.

[3] C. Fruh and A. Zakhor. An automated method for large-scale, ground-based city model acquisition. *International Journal of Computer Vision*, 60:5–24, 2004.

[4] T. Hackel, N. Savinov, L. Ladicky, J. D. Wegner, K. Schindler, and M. Pollefeys. SEMANTIC3D.NET: a New Large-Scale Point Cloud Classification Benchmark. *ISPRS*

[2] http://www.semantic3d.net/

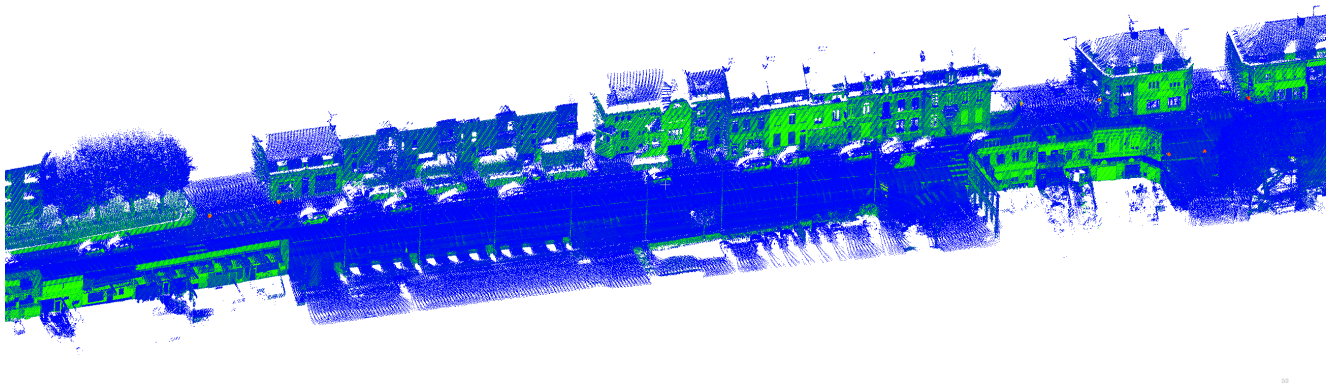

Figure 2. A section of the Paris-Lille-3D point cloud (reflectance ranges from 0 [blue] to 255 [red]).

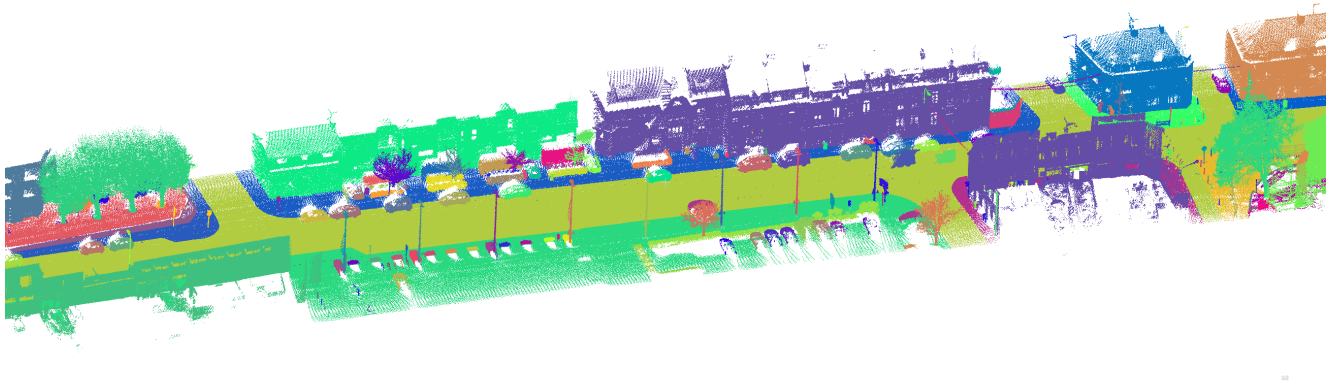

Figure 3. The manually labeled Paris-Lille-3D point cloud (one color was used for each object instance).

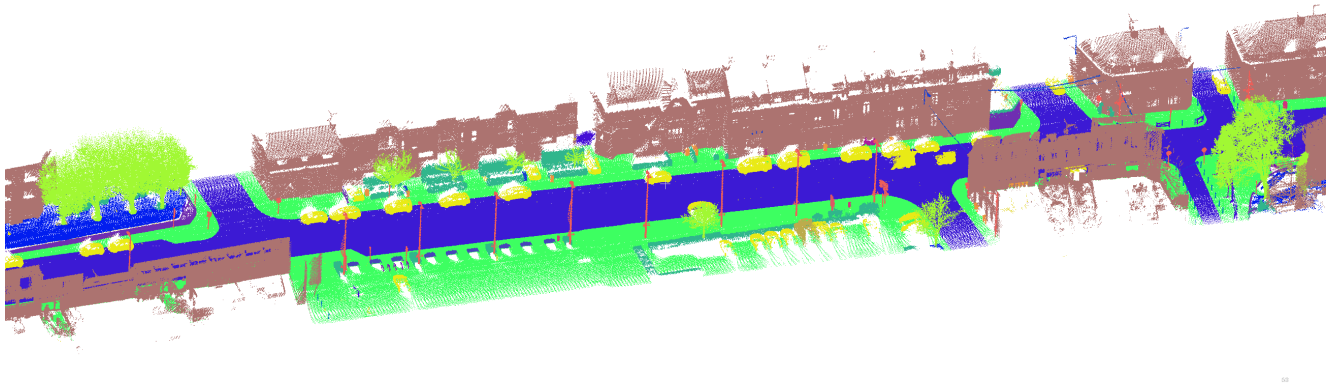

Figure 4. The classes of the Paris-Lille-3D point cloud (one color was used for each class).

Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, pages 91–98, May 2017.

[5] W. Maddern, G. Pascoe, and P. Newman. Leveraging experience for large-scale lidar localisation in changing cities. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1684–1691, May 2015.

[6] X. Roynard, J.-E. Deschaud, and F. Goulette. Paris-Lille-3D: a large and high-quality ground truth urban point cloud dataset for automatic segmentation and classification. *ArXiv e-prints*, Nov. 2017.

[7] X. Roynard, J.-E. Deschaud, and F. Goulette. Classification of Point Cloud Scenes with Multiscale Voxel Deep Network. *ArXiv e-prints*, Apr. 2018.

[8] A. Serna and B. Marcotegui. Detection, segmentation and classification of 3d urban objects using mathematical morphology and supervised learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93:243 – 255, 2014.

[9] B. Vallet, M. Brdif, A. Serna, B. Marcotegui, and N. Paparoditis. Terramobilita/iqmulus urban point cloud analysis benchmark. *Computers and Graphics*, 49:126 – 133, 2015.

[10] S. Wang, M. Bai, G. Mattyus, H. Chu, W. Luo, B. Yang, J. Liang, J. Cheverie, S. Fidler, and R. Urtasun. TorontoCity: Seeing the World with a Million Eyes. *ArXiv e-prints*, Dec. 2016.