SZTAKIBudapest: a multimodal Lidar benchmark for autonomous vehicles

August 16, 2021

Örkény Zováthi
research associate
SZTAKI

Balázs Nagy
research fellow
SZTAKI

Csaba Benedek
research advisor
SZTAKI

Contents

1 Introduction

In this report we present our collected multimodal *SZTAKIBudapest* Benchmark, a benchmark to carry both onboard Lidar vehicle measurement (OBM) data and mobile laser scanning (MLS) high density 3D city maps at the same places for the purpose of evaluating Lidar based point cloud segmentation, multimodal registration and change detection algorithms in urban environments.

2 Data acquisition

The onboard Lidar measurement sequences of the benchmark have been captured by SZTAKICar, a test car of our research institute in various main city roads of Budapest Hungary. The measurement platform has been equipped with a Velodyne HDL64E[1] 64-beam Rotating Multi-beam (RMB) Lidar scanner and a GPS receiver fixed on the roof top which provided rough global position estimations of the recorded point clouds.

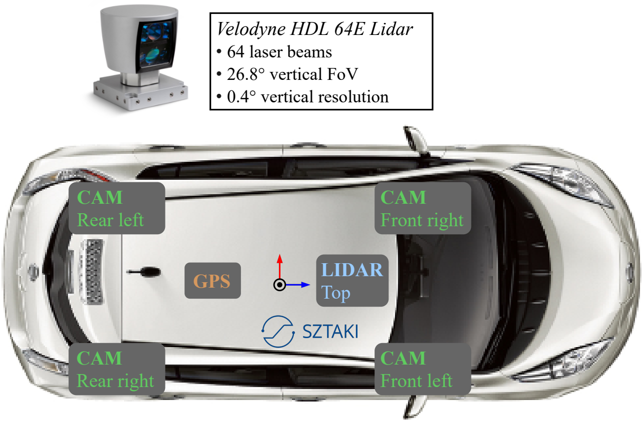


Figure 1: SZTAKICar: The measurement platform.

The Velodyne HDL-64E sensor was originally designed to help real-time perception of autonomous robots and vehicles. It provides a stream of relatively sparse (6-10 $\times$ 10^4 points/frame) point clouds with a temporal frequency of 15 fps. The spatial accuracy is around 1-2 cm in the sensor's own coordinate system, but the point density quickly decreases as a function of the distance from the sensor and it shows typical ring patterns.

The Mobile Laser Scanning (MLS) measurements have been recorded with a Riegl VMX450[2] mobile mapping system by the Budapest Road Management company (Budapest Közút Zrt.). The Riegl VMX450 MLS system is highly appropriate for city mapping, urban planning and road surveillance applications. It integrates two Riegl laser scanners, a well-designed, calibrated camera platform and a high performance Global Navigation Satellite System (GNSS), providing extremely dense, accurate (up to global

[1] http://velodyneLidar.com/
[2] http://www.riegl.com/

accuracy of a few centimetres) and feature rich data with a quite uniform point distribution.

3 Challenging tasks and datasets

In this section, we describe challenging point cloud segmentation, registration and change detection tasks on our Benchmark, and we provide two different datasets for evaluating evaluate such algorithms.

3.1 Automatic segmentation of MLS data (SZTAKI-CityMLS)

Introduction SZTAKI-CityMLS has been created for the purpose of evaluating 3D semantic point cloud segmentation algorithms in urban environments, based on mobile laser scanning (MLS) measurements of a Riegl VMX-450 mobile mapping system. Test data has been provided by Budapest Közút Zrt, an industrial partner of the research institute SZTAKI.

Figure 2: Sample point cloud data from urban environment. Raw point clouds (left), annotations (right). Annotation classes: **phantom**, **pedestrian**, **tram/bus**, **vehicle**, **vegetation**, **tall column** (including traffic sign holders and tree trunks), **street furniture** (various further street objects such as benches, dustbins, short columns), **ground**, **building facade**.

Motivation Recent deep learning based point cloud classification approaches such as the PointNet++ [2] or SPLATNet [3] show promising way to semantically segment various sorts on artificial or real laser scanning based point clouds. However, dealing with urban MLS data, a number of particular challenges appear – such as the phantom effect caused by independent object motions –, which are not handled by general point clouds segmentation algorithms efficiently enough. For this purpose, we created a new hand labeled dataset, called SZTAKI-CityMLS.

Technical details Our MLS dataset contains in total around 327 Million annotated points from challenging urban scenes in Budapest, including main roads with both heavy and solid traffic, public squares, parks, and sidewalk regions, various types of cars, trams and buses, several pedestrians and diverse vegetation. The annotation was manually constructed by a user friendly 3D point cloud annotator tool developed in our laboratory [1]. With this tool we manually labeled around 327M points over a 30.000 m^2 area of the city, with more than 50 m elevation differences, using the following nine classes: **phantom**, pedestrian, **tram/bus**, **vehicle**, **vegetation**, tall column (including traffic sign holders and tree trunks), **street furniture** (various further street objects such as benches, dustbins, short columns), **ground**, **building facade**. As ground truth (GT), we provide the same point clouds as the raw MLS measurements, assigned with color information to the corresponding class.

Access and contact Data and demo results [1] are available at the following link: `http://mplab.sztaki.hu/geocomp/SZTAKI-CityMLS-DB.html`
For more information, please contact Balázs Nagy (`nagy.balazs@sztaki.hu`, `balazs.nagy.it@gmail.com`) or Csaba Benedek (`benedek.csaba@sztaki.hu`).

3.2 Real-time Lidar only global localization and change detection (SZTAKI-CityCDLoc)

Introduction SZTAKI-CityCDLoc has been created for the purpose of evaluating multimodal 3D semantic point cloud registration and change detection algorithms in urban environments, based on mobile laser scanning (MLS) data of a Riegl VMX-450 mobile mapping system and onboard vehicle measurements captured by a Velodyne HDL 64E rotating multi-beam Lidar sensor.

Motivation In dense urban environment, we should expect that the initial position estimation of a vehicle might be notably inaccurate and the global positioning error of the vehicles may reach several meters in city regions with poor GPS signal coverage (Figure 4). Assuming that using an efficient segmentation [1] of the available raw MLS data, we can construct a subset of the segmented MLS point cloud, which contains static classes (tall column, street furniture, facade) and represents empty street segments. Therefore, we can consider it as highly detailed reference models for the vehicles' onboard Lidar measurements. In this context, accurate *global localization* of the vehicle's Lidar measurements in the MLS data, and the *detection of relevant changes* in the vehicle's environment based on the static MLS data may appear as challenging tasks.

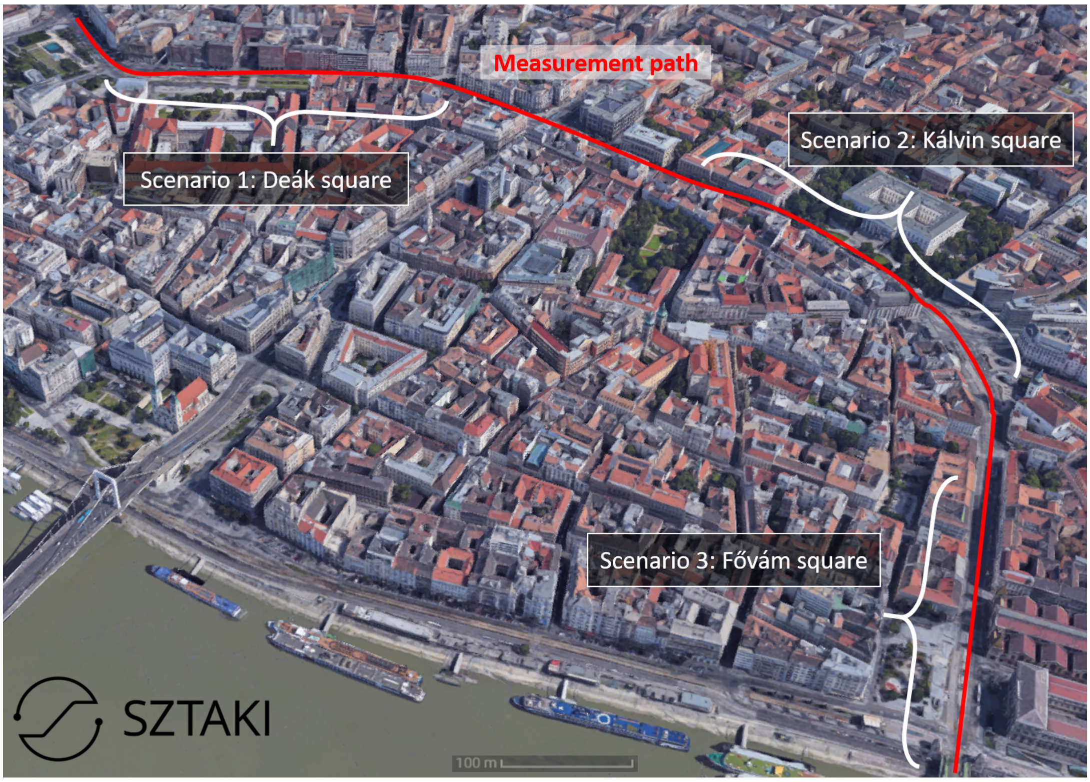


Figure 3: The data acquisition path and the three test scenarios

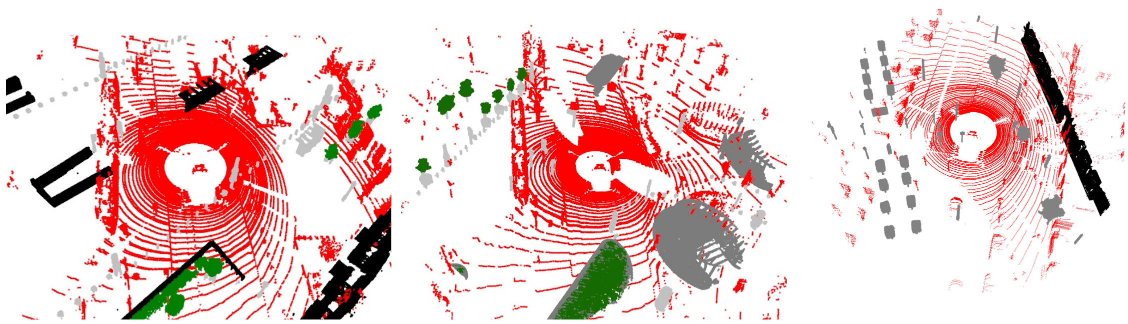
MLS scenario	Format	Start frame	End frame	Length [m]	Length [s]
Fővám square	#frame.pcd	8400	8600	300	13.33
Kálvin square	#frame.pcd	9500	9700	300	13.33
Deák square	#frame.pcd	11650	11730	200	5.33

Table 1: Lidar measurement sequences

Technical details We provide three test scenarios from the downtown of Budapest (Figure 3, Table 1). Each scenario contains the following data:

- a **geo-referred MLS point cloud** (*#scenario_MLS.pcd*) that is automatically segmented by [1] and it contains class labels regarding the facade, tall column and street furniture semantic classes (see Section 3.1).
- **RMB Lidar measurement sequences**. Each frame of the measurement is stored in a separate .pcd file (*#scenario_framenum_timestamp.pcd*). The coordinates of each frame are local to the sensor's center position. Beside the local XYZ coordinates, each point contains an intensity attribute as well.
- a separate **GPS metadata** (*gps.eov*), which contains timestamp-GPS coordinate pairs during the *whole measurement path* for positioning a Lidar frame in the geo-referred MLS data.

For the *global localization*, we do not provide any ground truth information, as there is no determinable "best" alignment between the significantly different point sets. Instead, we suggest to use error metrics to measure the efficiency of your alignment algorithms, such as the Modified Hausdorff Distance (MHD) or Median Point Distance (MPD).



(a) Scenario 1, Fővám square (b) Scenario 2, Kálvin square (c) Scenario 3, Deák square

Figure 4: GPS based positioning of the vehicles' Lidar measurements in the geo-referenced MLS data's coordinate system.

For *change detection*, we annotated uniformly sampled Lidar point cloud measurement frames from each test scenario, and labelled the ground truth information in a semi-automatic manner. First, we performed an approximate offline registration between the i3D and MLS frames using the Iterative Closest Point (ICP) [4] algorithm, then we applied an automated nearest neighbor search based classification with a small distance threshold (5 cm) as an initial segmentation result. Thereafter, the labeling of the different change regions (especially on the region borders) in the sampled Lidar frames was manually revised using our previously mentioned 3D point cloud annotator tool [1]. At the annotation, we distinguished four change classes (Figure 5) by GT labeling:

- *Dynamic changes* that refer either to moving street objects such as traffic participants, to temporarily available objects such as barriers, or to changes in static scene elements such as a re-located bus station or kiosk. These regions are not presented in the MLS data. At GT labeling, these points are marked with **red** ($r = 255$, $g = 0$, $b = 0$).

- *Seasonal changes*, which regions are typical for vegetation areas. These regions are segmented as vegetation in the MLS data, and may have modified appearance during the different time periods/seasons. At GT labeling, these points are marked with **dark green** ($r = 0$, $g = 128$, $b = 0$).

- *Unchanged regions*, which contain static environment parts. These regions are also present in the MLS data. At GT labeling, these points are marked with **blue** ($r = 0$, $g = 0$, $b = 255$).

The annotations are stored in a separate folder in .pcd files with the same name as the original measurement frame, where the given colors code the corresponding class for each point.

(a) Scenario 1, Fővám square (b) Scenario 2, Kálvin square (c) Scenario 3, Deák square

Figure 5: Change detection GT labeling samples from each test scenario. Color codes: dynamic change, seasonal change, no change.

Access and contact For accessing this dataset, please send an e-mail request to Örkény Zováthi (`zovathi.orkeny@sztaki.hu`) or to Csaba Benedek (`benedek.csaba@sztaki.hu`).

References

[1] B. Nagy and C. Benedek. 3D CNN-based semantic labeling approach for mobile laser scanning data. *IEEE Sensors Journal*, 19(21):10034–10045, Nov 2019.

[2] C.R. Qi, L. Yi, H. Su, and L.J. Guibas. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. In *Conference on Neural Information Processing Systems (NIPS)*, pages 5105–5114, Long Beach, CA, USA, 2017.

[3] Hang Su, Varun Jampani, Deqing Sun, Subhransu Maji, Evangelos Kalogerakis, Ming-Hsuan Yang, and Jan Kautz. SPLATNet: Sparse lattice networks for point cloud processing. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2530–2539, 2018.

[4] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.