


Multi-sensorial Environment Perception in Urban Environment

ROBOVIS 2020 – Keynote talk

Csaba Benedek

Machine Perception Research Laboratory, Institute for Computer Science and Control (SZTAKI), Hungary

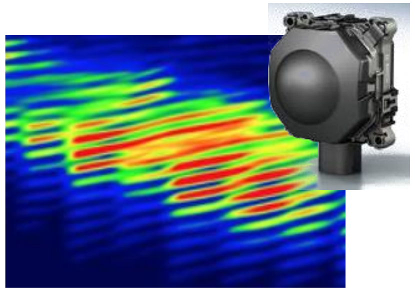
benedek.csaba@sztaki.hu

Research background

- Institute for Computer Science and Control (SZTAKI),
  - most prestigious IT research institute of Hungary, which bridges the gap between academic research and industrial expectations.
- SZTAKI Machine Perception Research Laboratory (MPLab), founded in 2006
  - Goal: interpretation and organization of information coming from distributed multimodal sensors

High resolution cameras

Thermal cameras

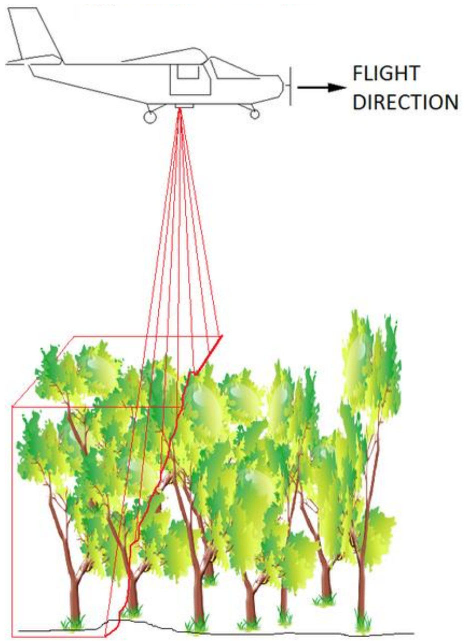


Radar

LIDAR laser scanners

- Research Group on Geo-information Computing (GeoComp) @ MPLab headed by Csaba Benedek
  - Research objectives: large scale spatial data management and filtering, automated 3D scene understanding used in various applications
  - Applied methodologies: Computer vision, geometry based and probabilistic modeling, machine learning applications

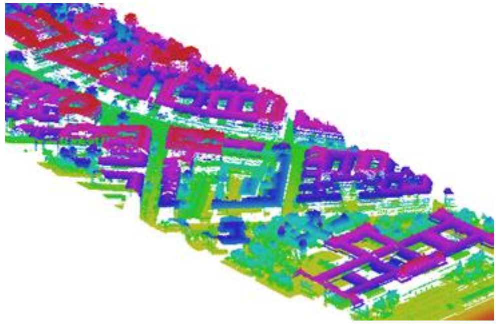
LIDAR technology



LIDAR = LIght Detection and Ranging

- surveying method that measures distance to a target by illuminating it with laser light and measuring the reflected light with a sensor (≈ ***laser radar)***
- rapidly provides large amount of accurate range data from 100+ meter distances

LIDAR platforms:

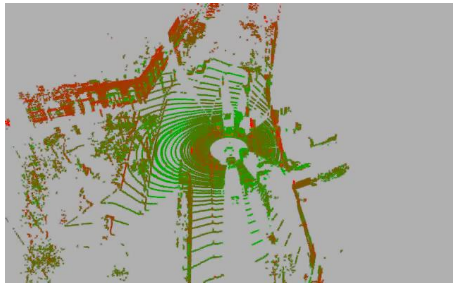


Aerial LIDAR

BUDAPEST KÖZÚT

Ground based mobile mapping system

Terrestrial static laser scanner



Real time mobile LIDAR sensors

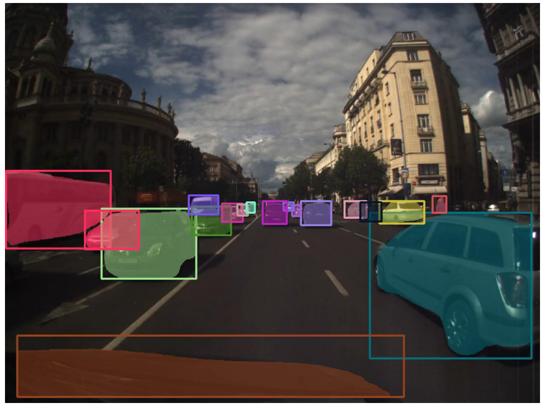
Instant environment perception from a mobile platform with a new generation geospatial database background

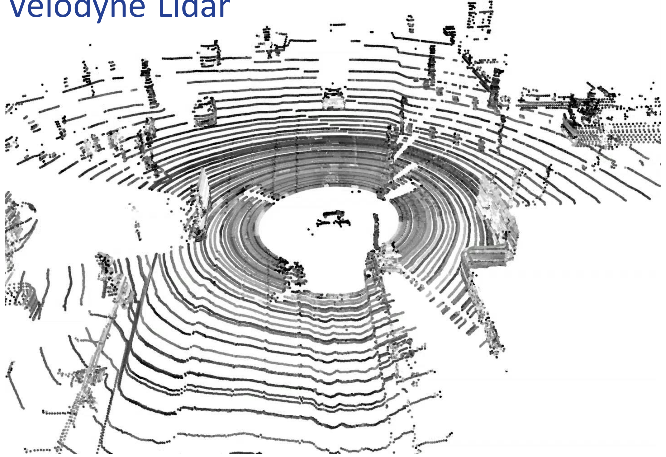
Project of the Hungarian National Research, Development and Innovation Fund (2016-2021)

- Car-mounted multisensory platform

Cameras

Real time Velodyne Lidar

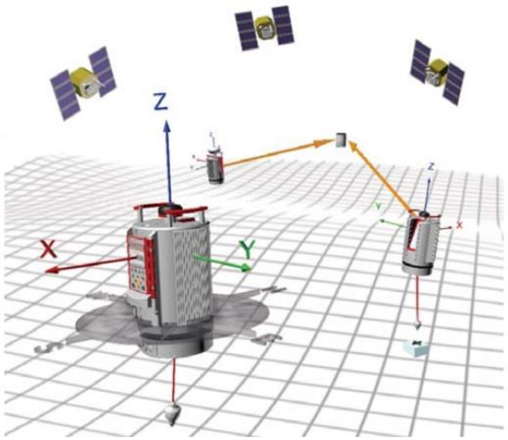




- High resolution localization map

- Accurate environment perception and modeling in smart cities based on multiple sensors
- Challenges:
  - Different sensor types and different dimension of the data
  - Synchronization issues
  - Accurate Lidar - Lidar and Lidar - Camera calibration is essential

Environment modeling: static and mobile laser scanning data



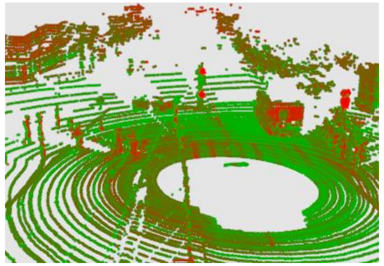
DSR camera

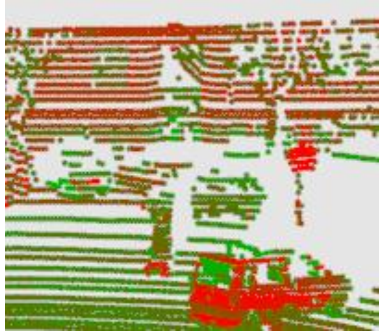
Scanner

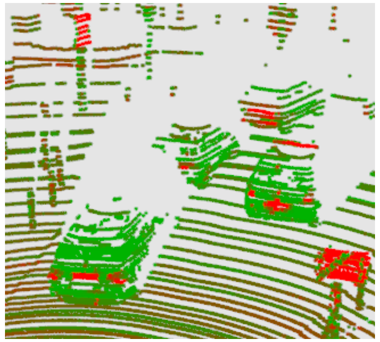
Output: colored point cloud

Pros and issues

i3D data





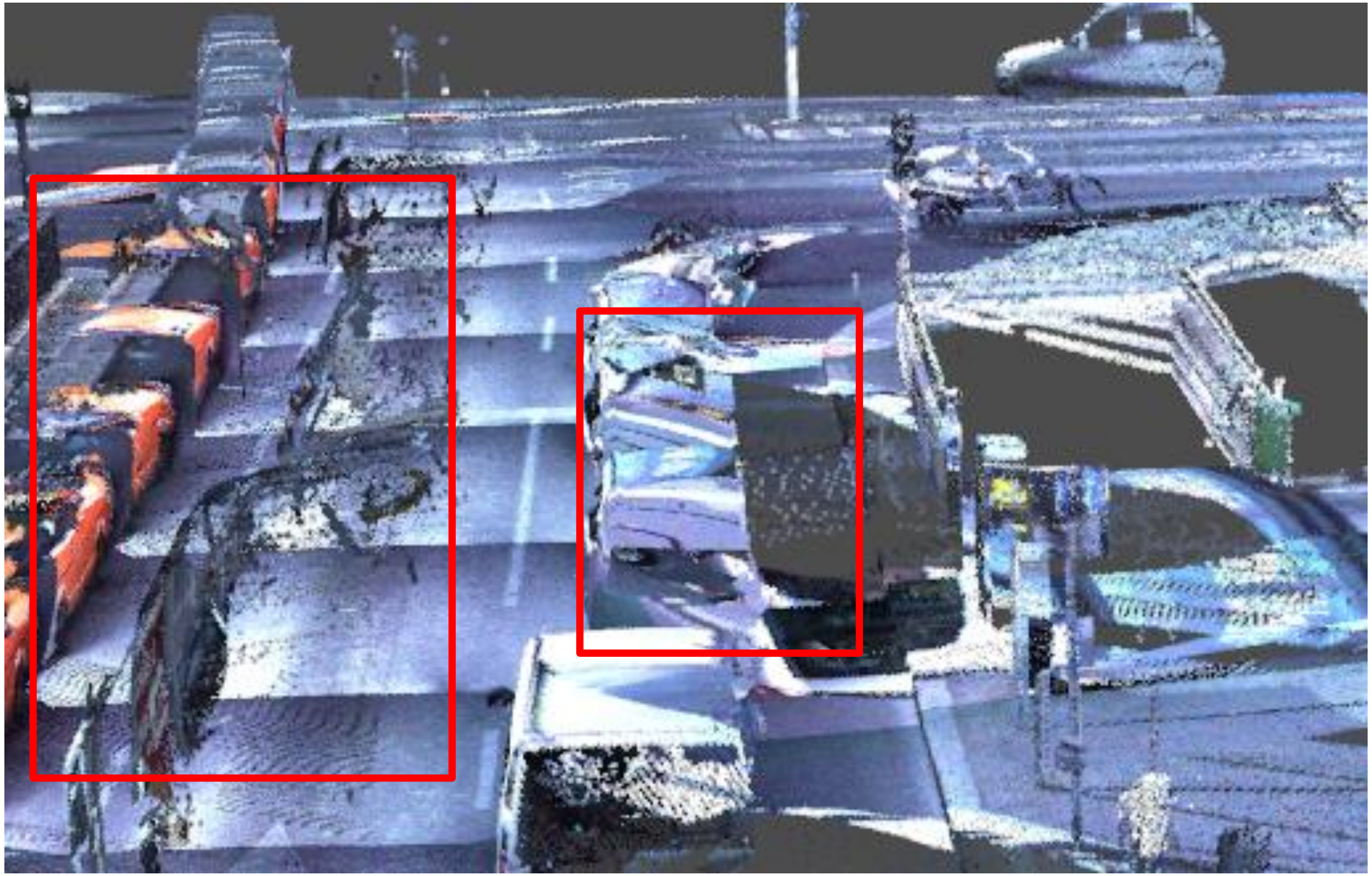


- Quite sparse, less accurate, only coarse geo-coordinates, strongly inhomogeneous point clouds (typical ring patterns)
- Instant data: situation analysis is possible, moving street objects can be observed and analyzed
- Online processing required
- Region classification, object separation and recognition is very challenging due to the particular point cloud characteristics
- Compact point cloud frames (max 60-100k points) – easy management and data transfer

MLS data

- Very dense, very accurate, geo-refered, quite homogeneous point clouds (linear density characteristics wrt. sensor dist),
- Only static environment model
- Offline processing allowed
- Ghosts of moving objects
- Parking cars, standing pedestrians, temporary street furniture, vegetation change
- Huge data: computational issues, memory issues, data transfer issues

Phantom object effects by mobile laser scanning

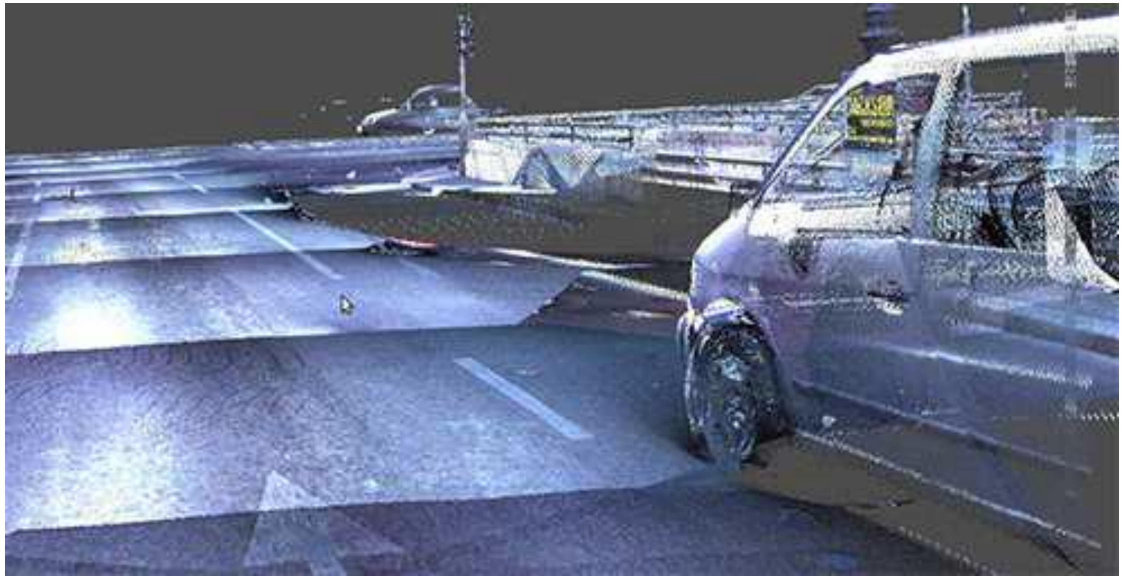


RIEGL VMX-450

B. Nagy, and *Cs. Benedek*: "3D CNN Based Phantom Object Removing from Mobile Laser Scanning Data", ***International Joint Conference on Neural Networks (IJCNN),*** pp. 4429-4435, Anchorage, Alaska, USA, 14-19 May, 2017

Phantom object effects by mobile laser scanning

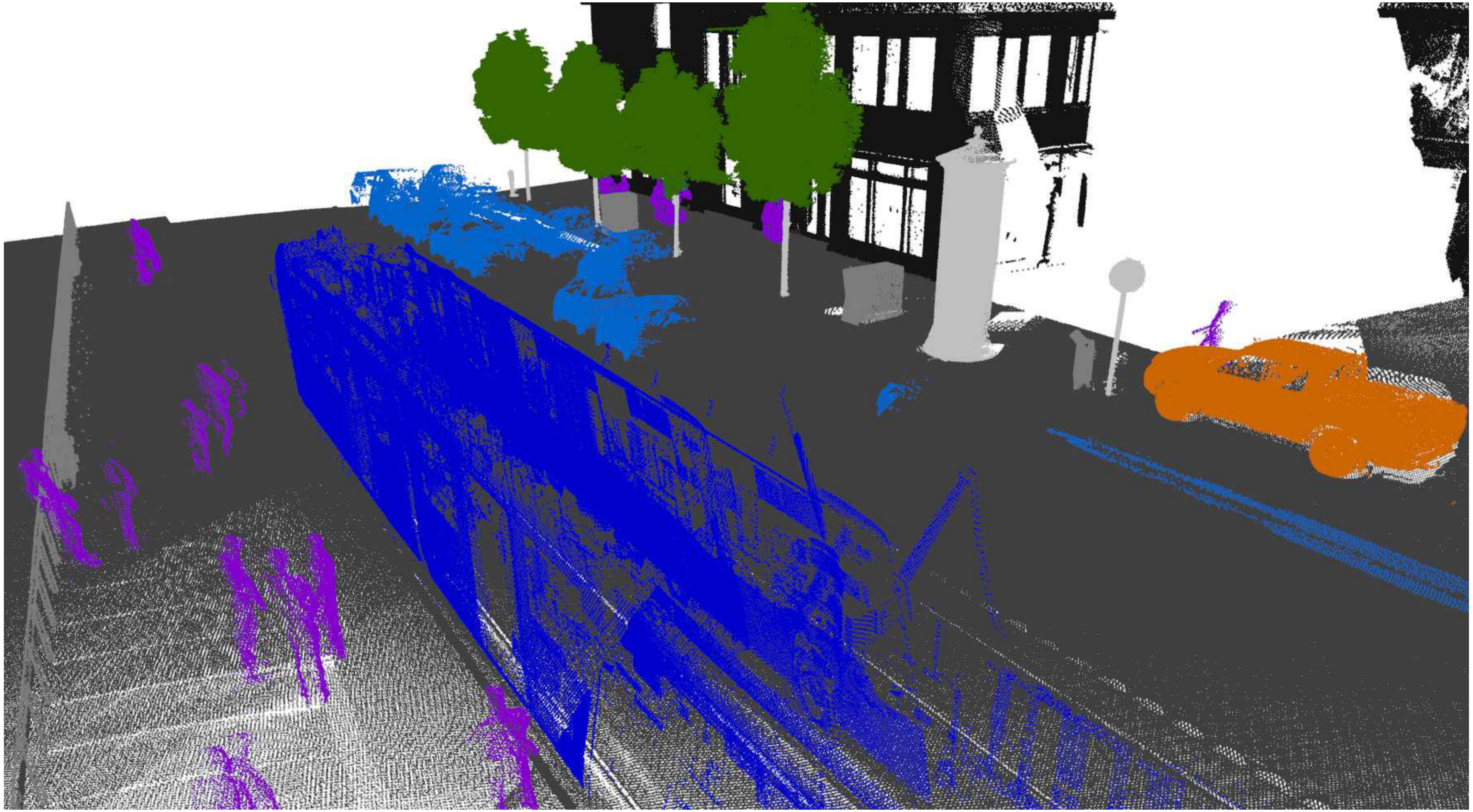
- 3D point cloud representation: detected phantoms can be removed



Phantom region detection

- Challenges of phantom region detection
  - Usually sparser than static objects, but…
  - MLS clouds have considerably inhomogeneous point density due
    - occlusions
    - different distances of the surface points from the scanner's trajectory
    - varying speed of the scanner platform (e.g. accelerating and breaking)
    - different laser adsorption/reflection properties of the different surface materials
  - Highly diverse phantom regions
    - varying speed of moving objects
- Proposed solution: deep neural networks – 3D semantic segmentation task
  - Apart from phantom detection let us distinguish different classes

Goal: semantic point cloud segmentation with 9 classes



- pedestrian
- phantom
- parking vehicle
- tram/bus
- column/tree trunk
- vegetation
- street furniture
- road
- building facade

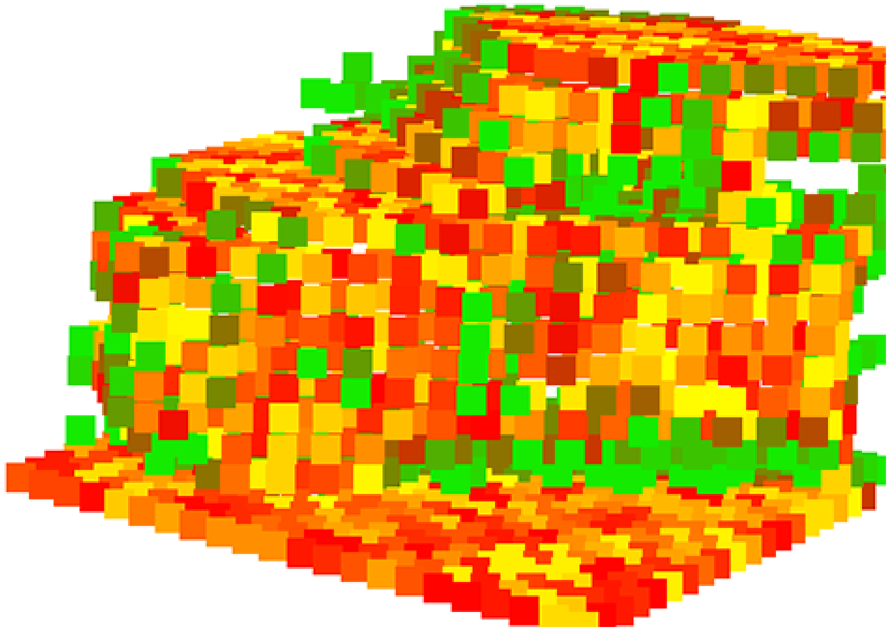
Existing point cloud classification solutions

- *Global* approaches
  - information from the complete 3D scene for classification of the individual voxels - main challenge: time and memory requirements
    - OctNet (expensive training data annotation)
- Sliding volume based techniques
  - move a 3D box over the scene, using locally available information for the classification of each point cloud segment.
    - *Vote3Deep* (fixed object size)
    - OG-CNN (purely local features)
    - multi-view technique (projection + 2D CNN models)
    - Latest: PointNet++ (indoors), SGPN (limited scene size), SPLATNet3D (discrete grid)
- Proposed solution: sliding volume based technique, but using both local and global features in data representation

Data representation

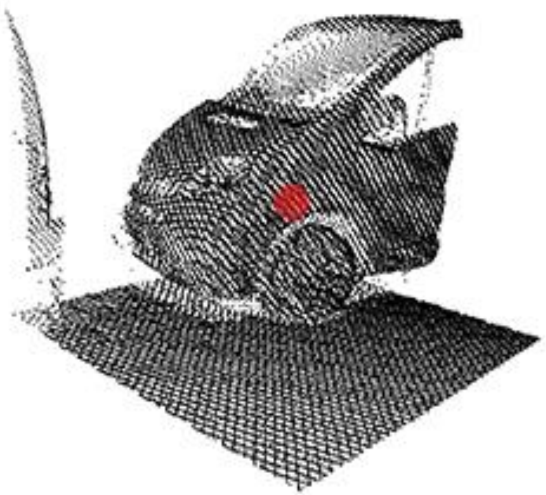
- Sparse voxel structure for the input point cloud, with a fine resolution (used 0.1m voxel side length)

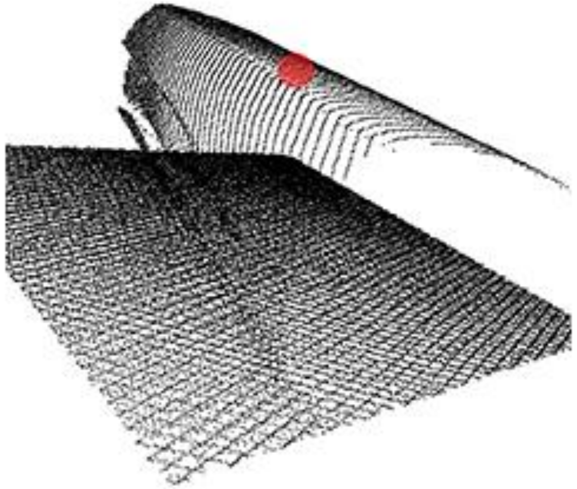
- Two feature channels:
  - point *density:* number of included points,
  - *mean elevation*, average of the point height values
- Training volume:
  - K × K × K voxel neighborhood (used K = 23),
  - unit of training and recognition in our network
  - Central voxel is classified with considering the whole training volume

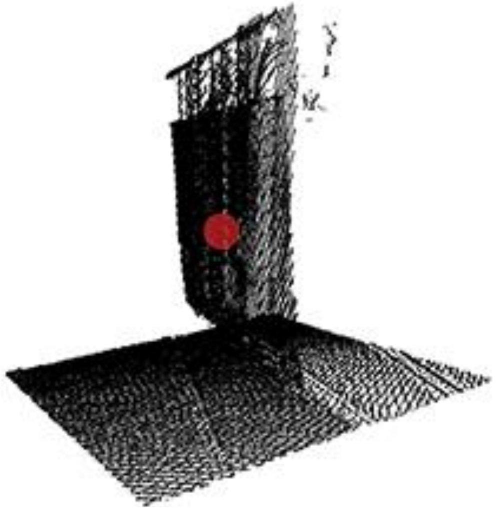


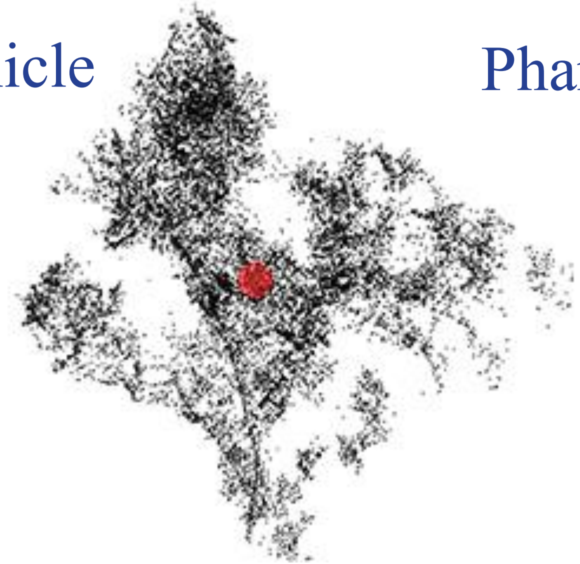
Voxelized training sample

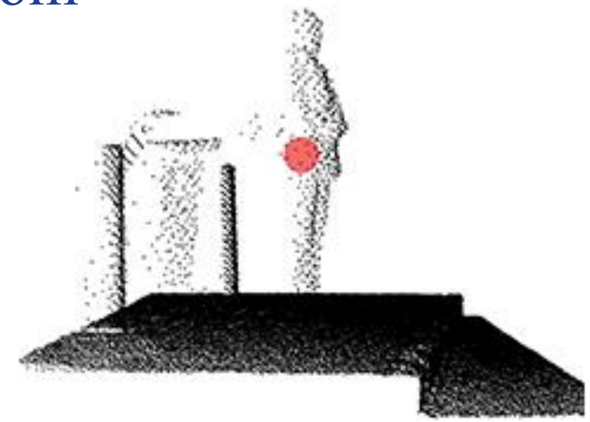
Training data representation and dataset generation


Vehicle

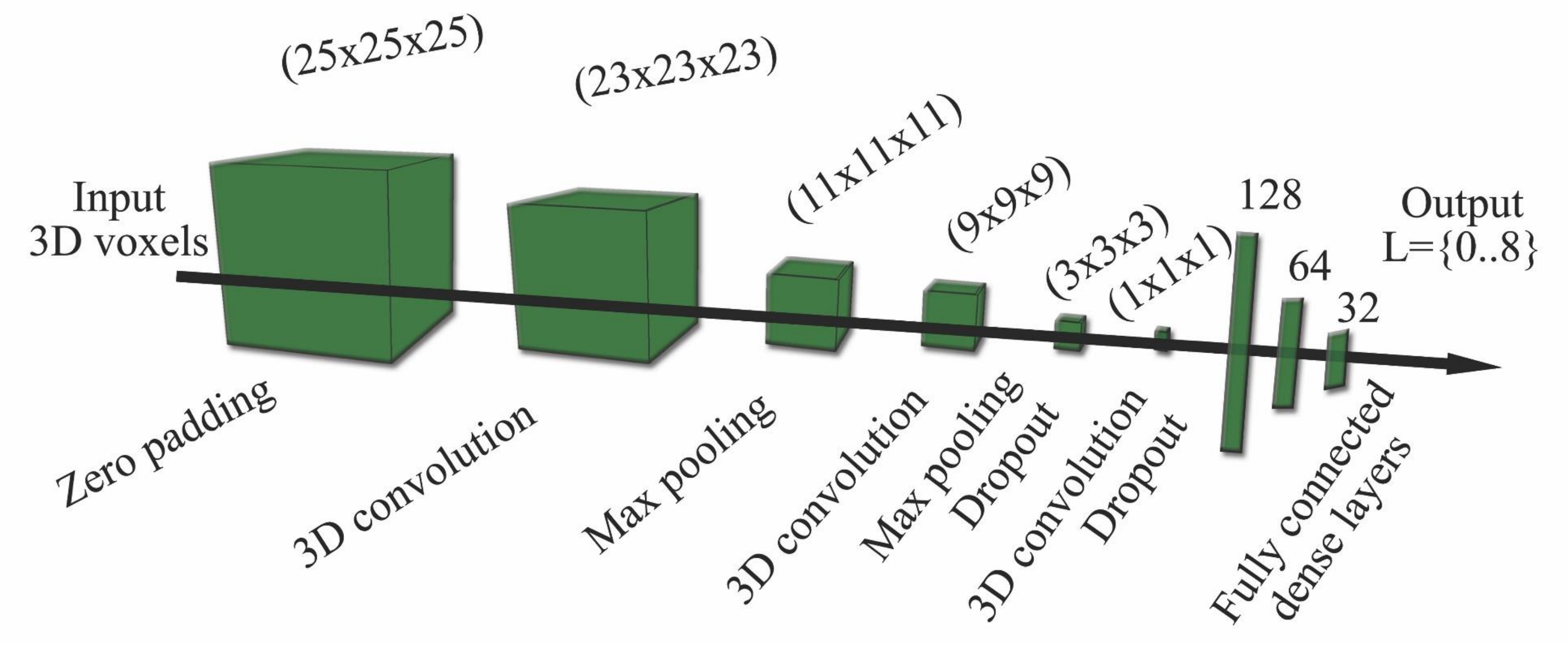

Phantom


Wall


Vegetation


Pedestrian

3D CNN



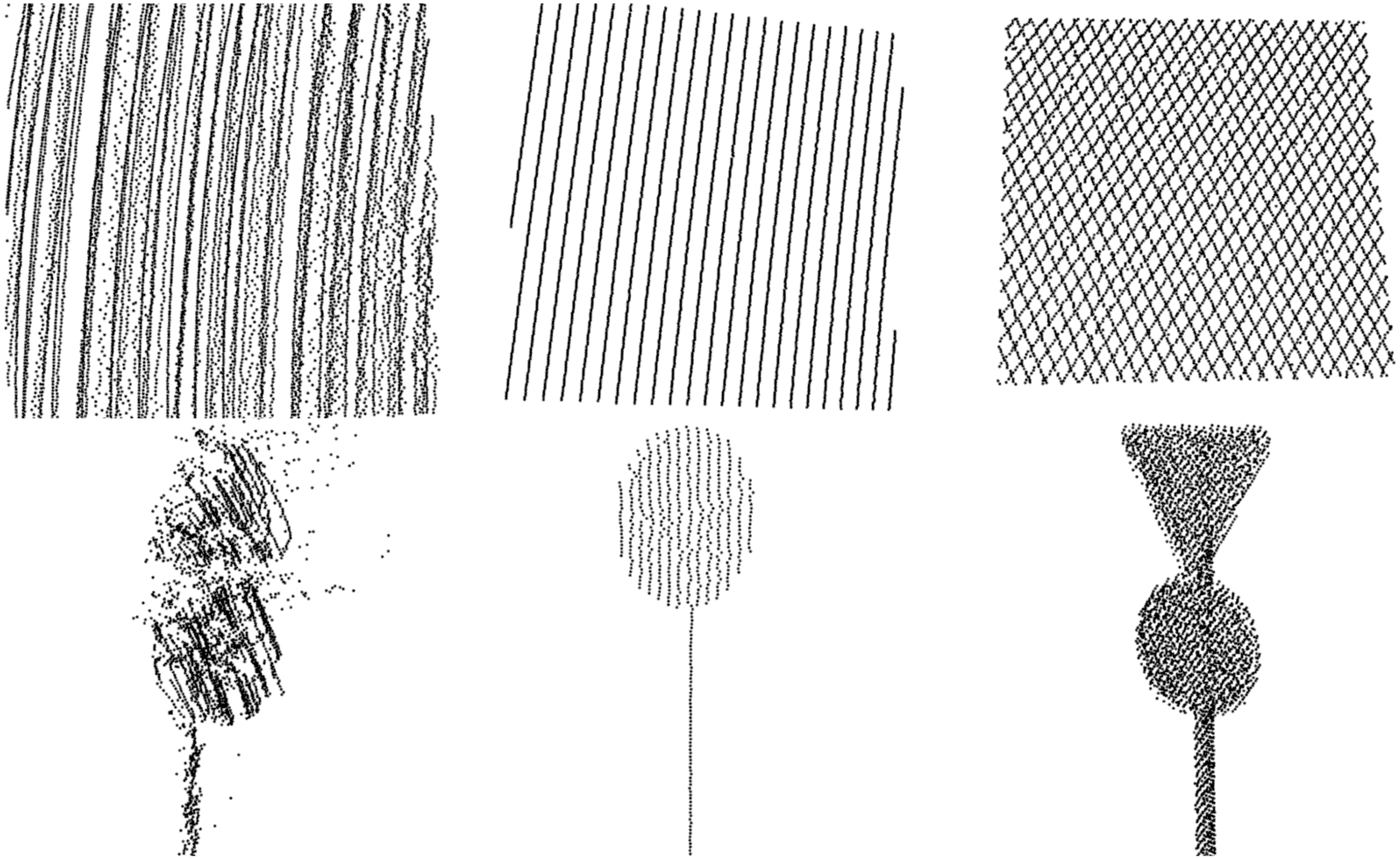
3D CNN architecture

Benchmarking issues

- Existing Mobile laser scanning (MLS) point cloud datasets - relatively small annotated segments
  - Oakland (1.6M points),
  - Paris-rue-Madame (20M points)
  - IQmulus & TerraMobilita (12M labeled points)
- Proposed dataset: **SZTAKI-CityMLS**
  - 327 Million annotated points
  - various urban scenes, 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.

URL: http://mplab.sztaki.hu/geocomp/SZTAKI-CityMLS-DB.html

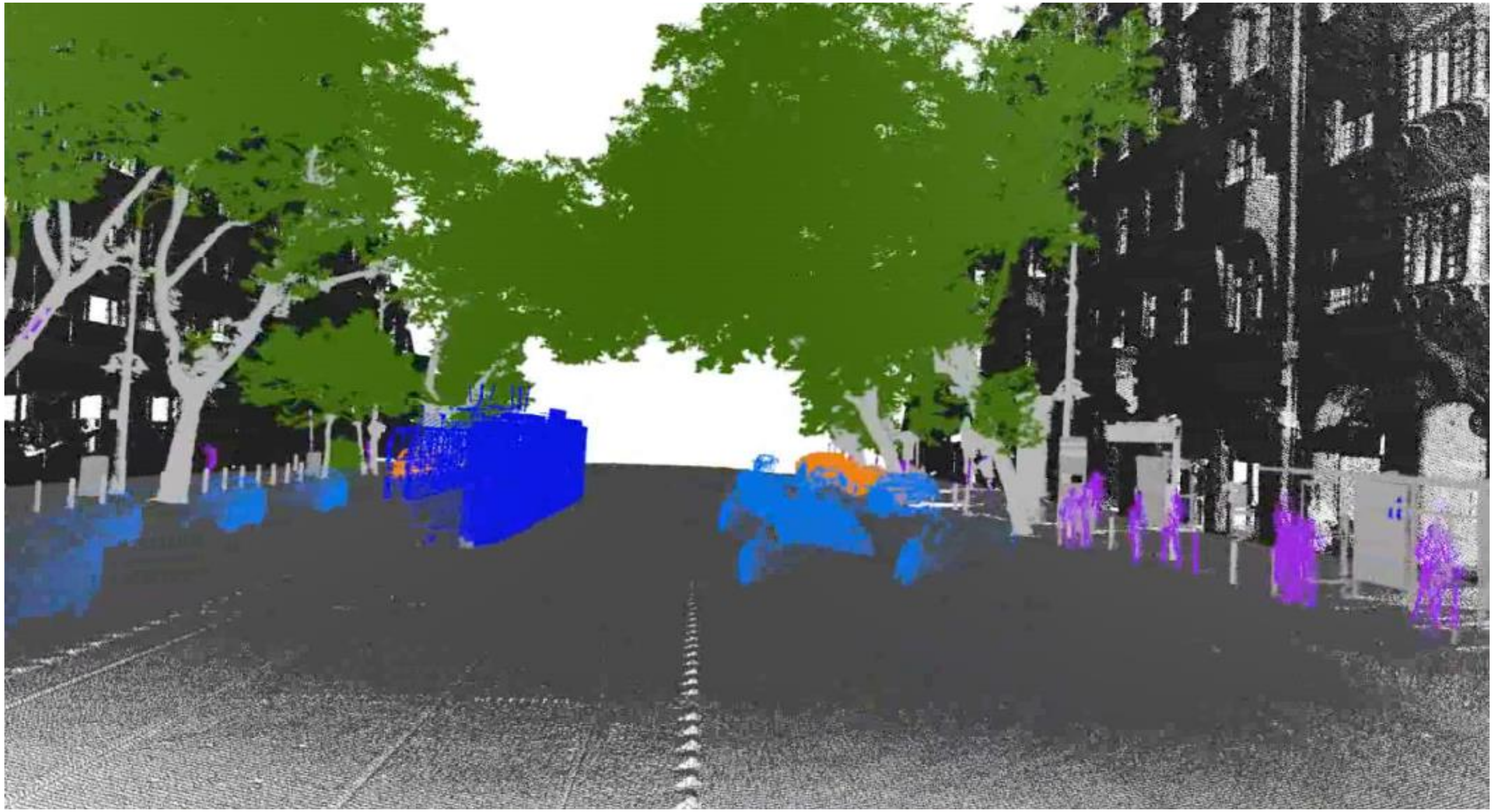
Data characteristic comparison



Paris-rue-Madame TerraMobilita SZTAKI-CityMLS

Semantic segmentation of MLS point clouds

- ***Offline*** utilization: mapping, road state estimation



pedestrian

moving vehicle

parking vehicle

tram/bus

column/tree trunk

vegetation

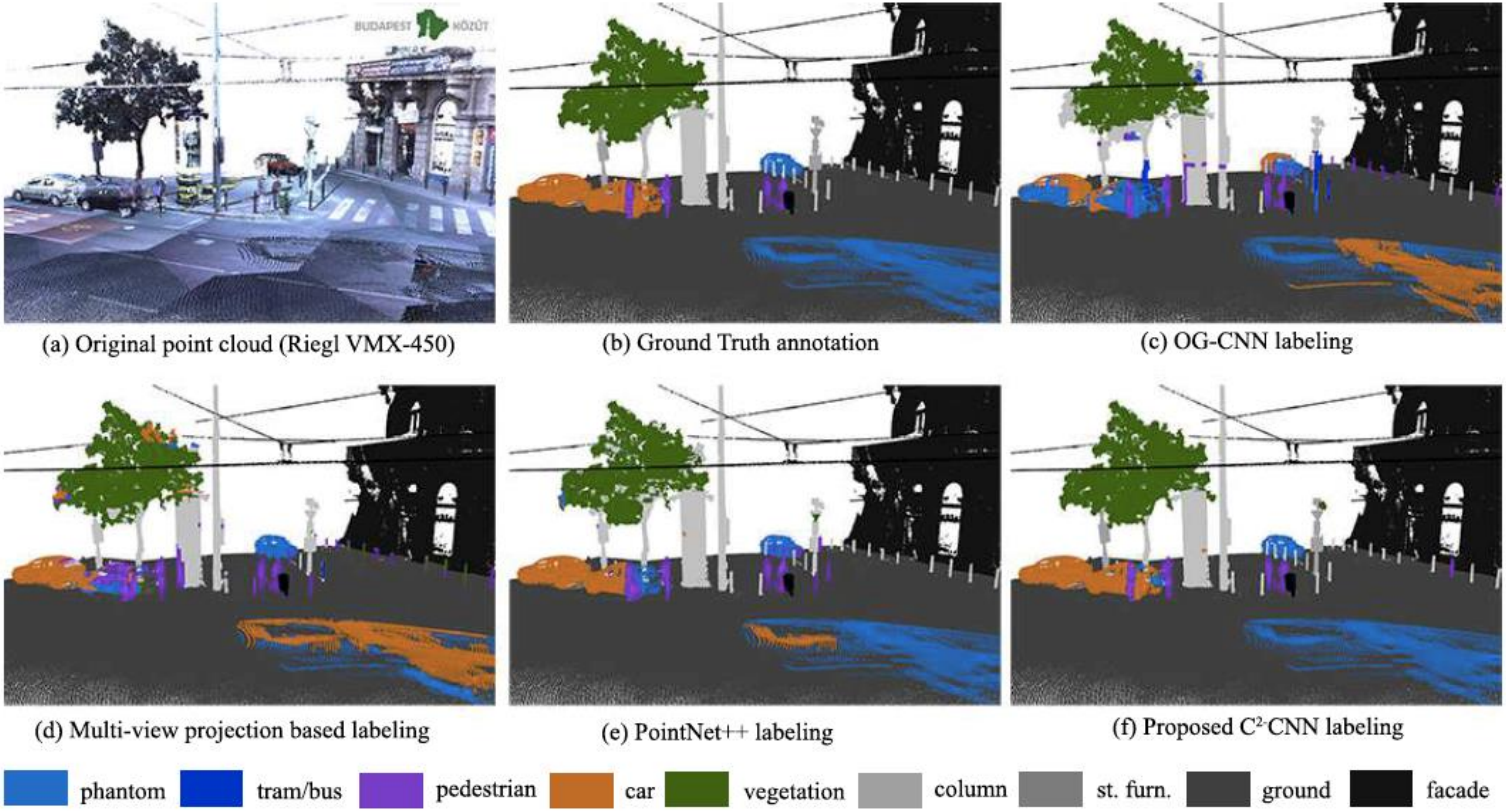
road

building facade

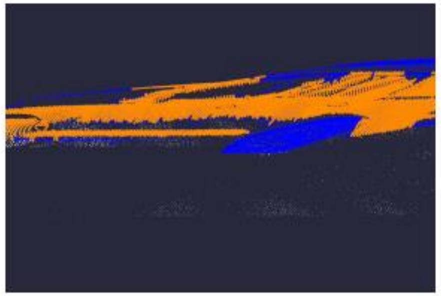
B. Nagy, and Cs. Benedek: "3D CNN Based Semantic Labeling Approach for Mobile Laser Scanning Data", ***IEEE Sensors Journal***, vol. 19, no. 21, pp. 10034 – 10045, 2019

BUDAPEST KÖZÚT

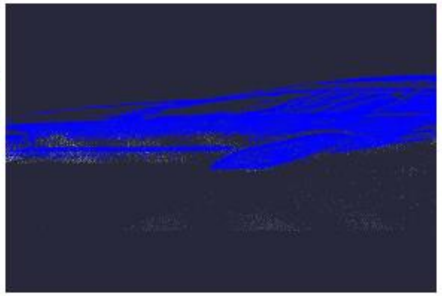
Evaluation


BUDAPEST KÖZÚT
(a) Original point cloud (Riegl VMX-450)
(b) Ground Truth annotation
(c) OG-CNN labeling
(d) Multi-view projection based labeling
(e) PointNet++ labeling
(f) Proposed C²-CNN labeling
phantom
tram/bus
pedestrian
car
vegetation
column
st. furn.
ground
facade

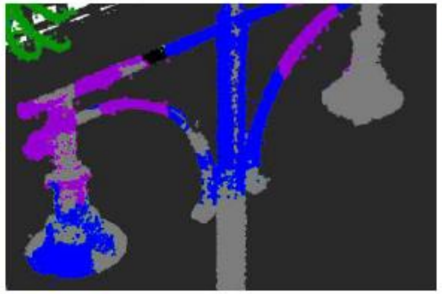
Comparison to reference methods



(a) PointNet++



(b) Proposed C^{2}CNN method



(c) PointNet++

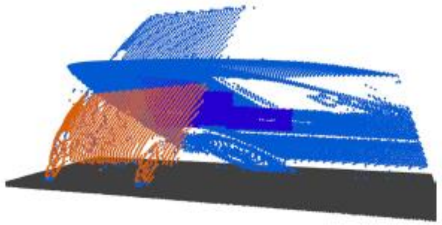
(d) Proposed C^{2}CNN method

(e) PointNet++

(f) Proposed C^{2}CNN method

(g) SPLATNet$_{3D}$

(h) Proposed C^{2}CNN method



(i) SPLATNet$_{3D}$

(j) Proposed C^{2}CNN method

(k) SPLATNet$_{3D}$

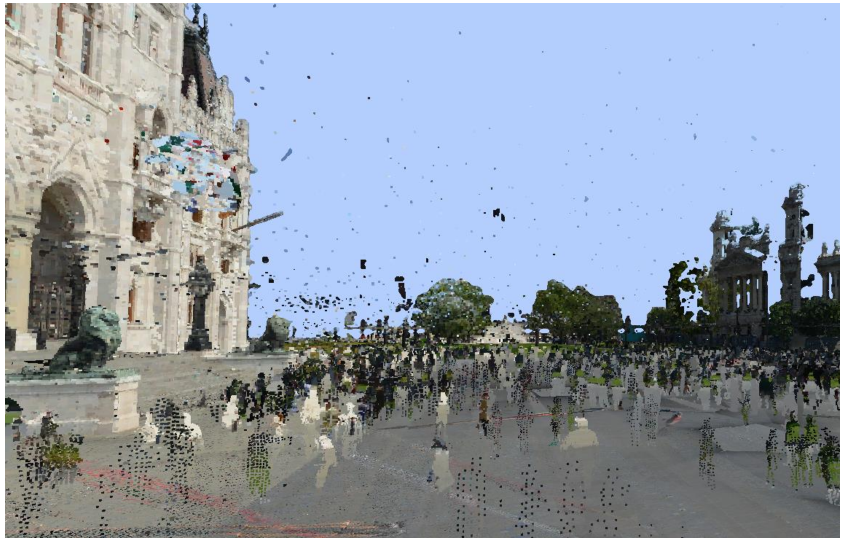
(l) Proposed C^{2}CNN method

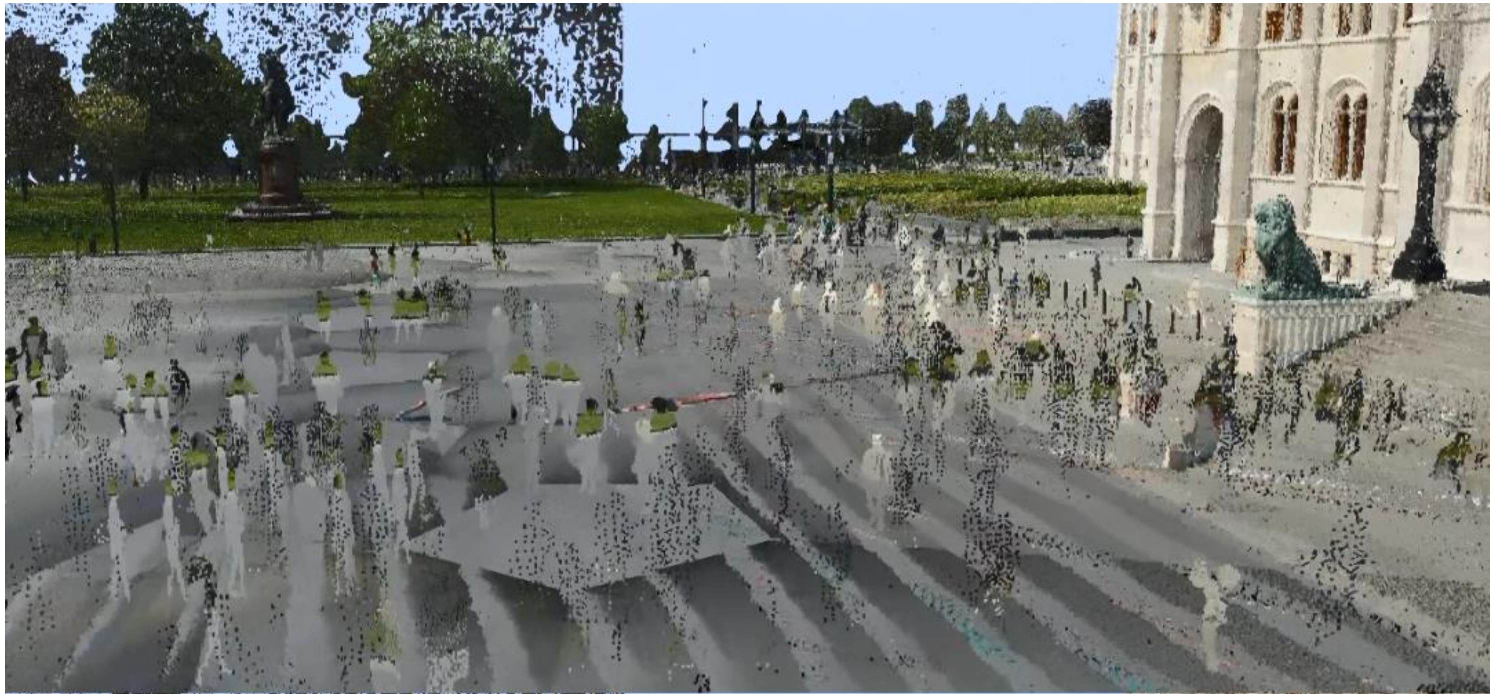
Comparison to reference methods

Class	OG-CNN [15]			Multi-view [18]			PointNet++ [7]			SPLATNet^{xyz} [20]			SPLATNet$^{xyz}_{rgb}$ [20]			Proposed C^{2}CNN		
	Pr	Rc	F-r	Pr	Rc	F-r	Pr	Rc	F-r	Pr	Rc	F-r	Pr	Rc	F-r	Pr	Rc	F-r
Phantom	85.3	34.7	49.3	76.5	45.3	56.9	82.3	76.5	79.3	82.5	80.9	81.7	83.4	78.2	80.7	84.3	85.9	**85.1**
Pedestrian	61.2	82.4	70.2	57.2	66.8	61.6	86.1	81.2	83.6	82.6	82.1	82.3	80.4	78.6	79.5	85.2	85.3	**85.2**
Car	56.4	89.5	69.2	60.2	73.3	66.1	80.6	92.7	86.2	81.5	90.0	85.5	81.1	89.4	85.0	86.4	88.7	**87.5**
Vegetation	72.4	83.4	77.5	71.7	78.4	74.9	91.4	89.7	90.5	87.1	88.2	87.6	86.4	87.3	86.8	98.2	95.5	**96.8**
Column	88.6	74.3	80.8	83.4	76.8	80.0	83.4	93.6	**88.2**	84.3	90.2	87.2	84.1	89.2	86.6	86.5	89.2	87.8
Tram/Bus	91.4	81.6	86.2	85.7	83.2	84.4	83.1	89.7	86.3	82.1	83.5	82.8	79.3	82.1	80.7	89.5	96.9	**93.0**
Furniture	72.1	82.4	76.9	57.2	89.3	69.7	84.8	82.9	83.8	84.7	86.2	**85.4**	82.6	81.3	81.9	88.8	78.8	83.5
Overall	76.9	74.2	75.5	72.5	73.4	72.9	85.6	87.5	86.5	83.5	85.9	84.7	82.5	83.7	83.0	90.4	90.2	**90.3**

Note: Voxel level Precision (Pr), Recall (Rc) and F-rates (F-r) are given in percent (overall values weighted with class significance)

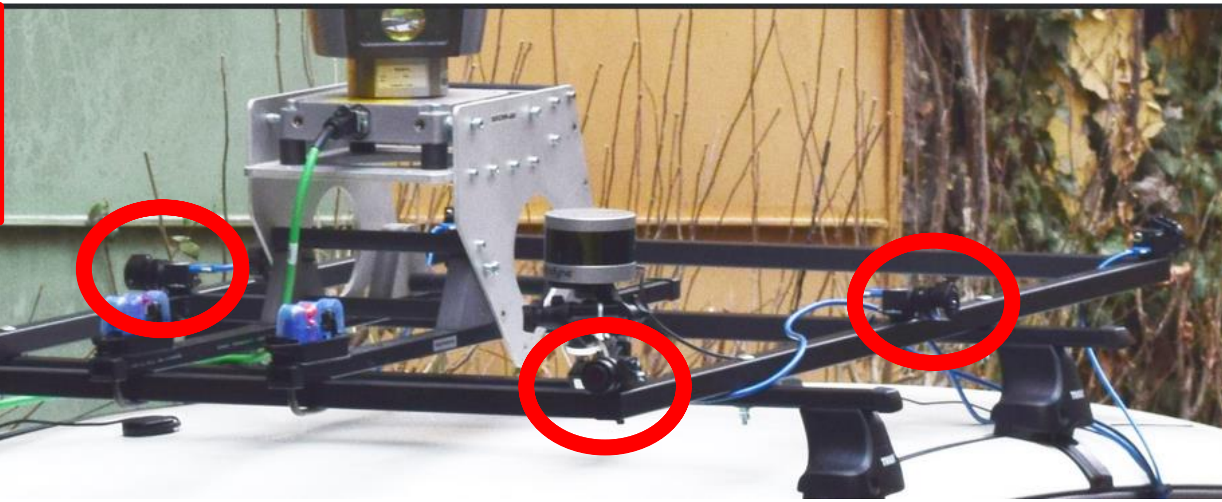
Cleaning of terrestrial laser scanning data

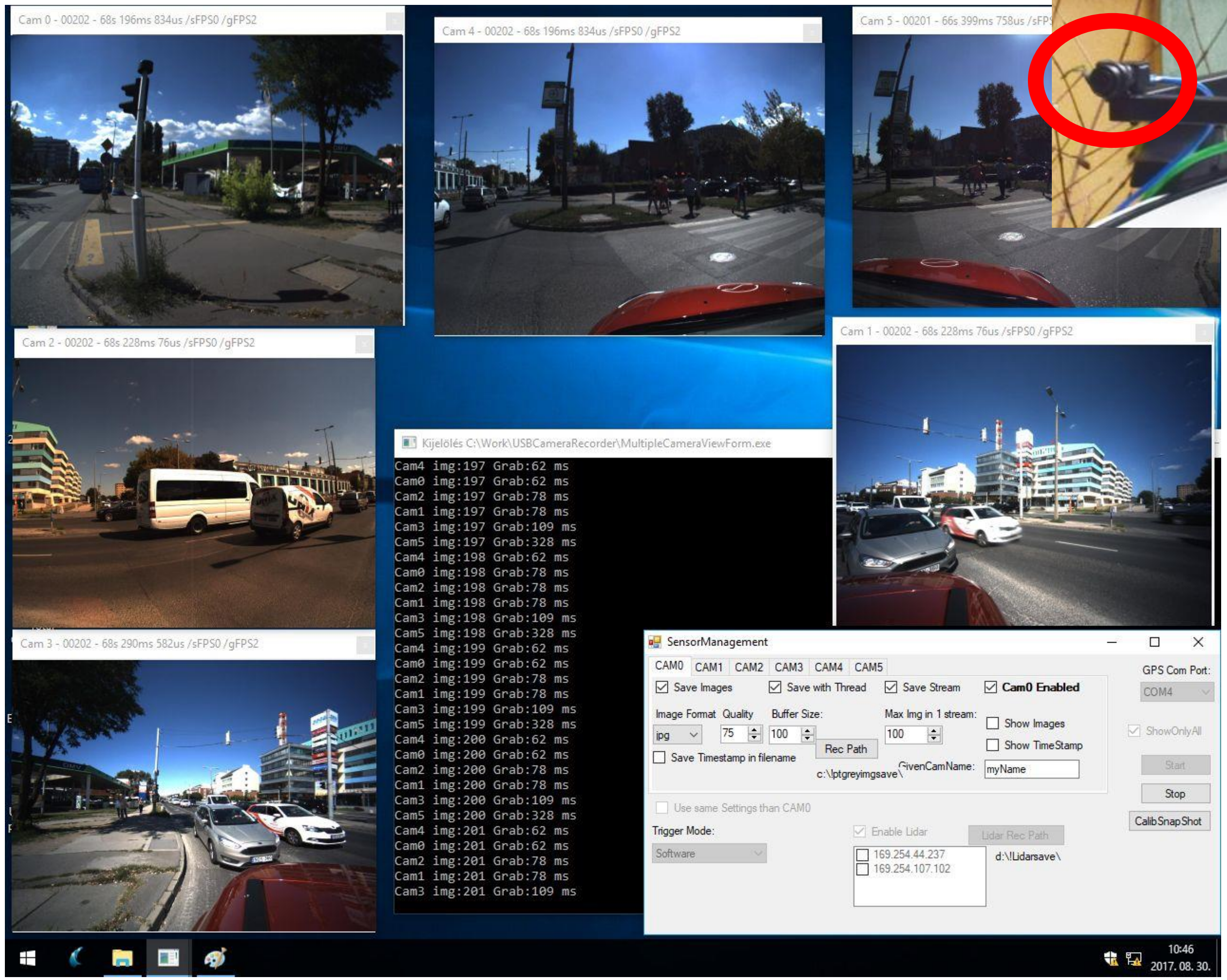




MTA SZTAKI BUDAPEST KÖZÚ

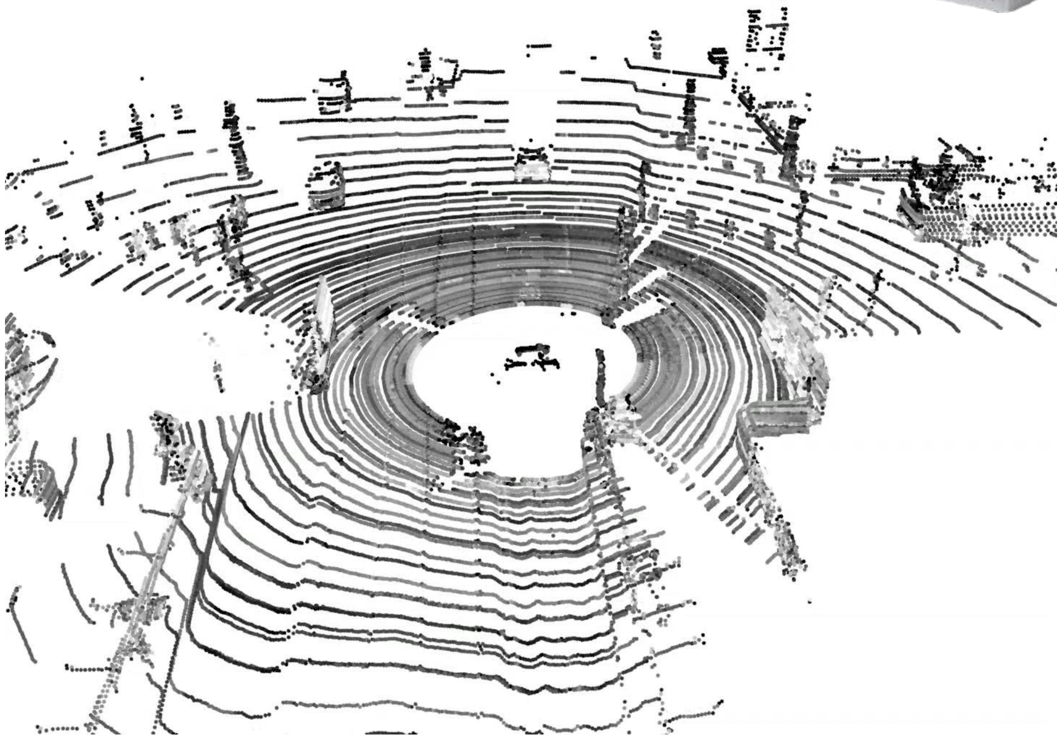
Lidar-camera platform





P. Gáspár, T. Szirányi, L. Hajder, A. Soumelidis, Z. Fazekas and C. Benedek. "Adding Autonomous Features to a Production Electric Car." **ERCIM News** 2017

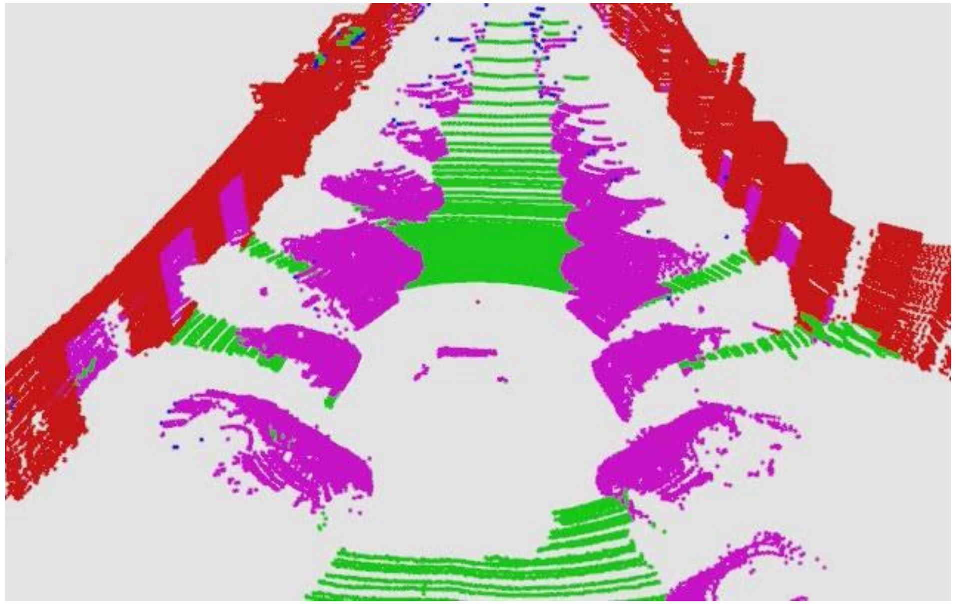
Rotating multi-beam Lidar data acquisition

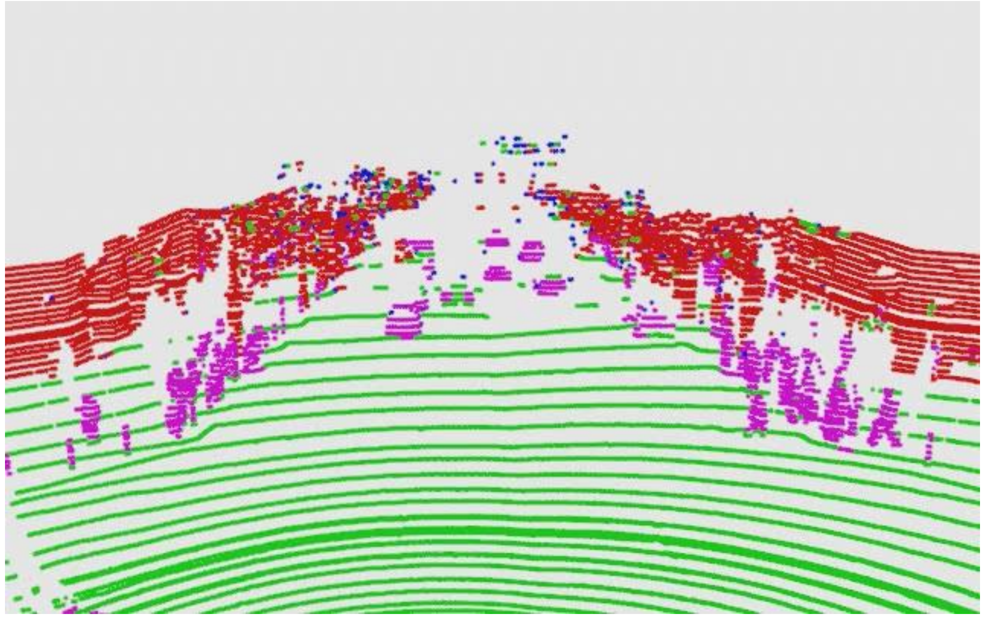


Ground-obstacle separation for i3D data flow

- 2D grid based locally adaptive terrain modeling
  - expect inhomogeneous RMB Lidar point clouds with typically non-planar ground.

Point cloud classification: Road Wall/lamp post Street objects



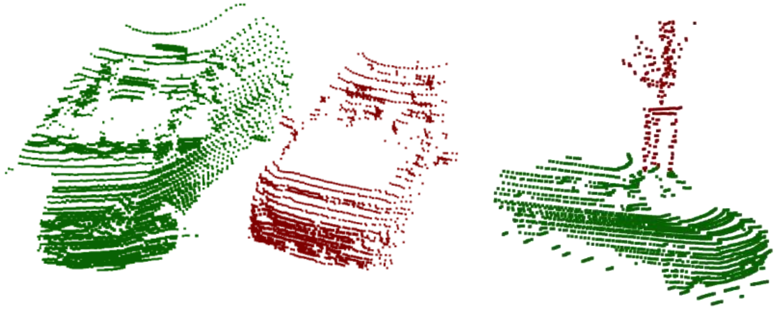


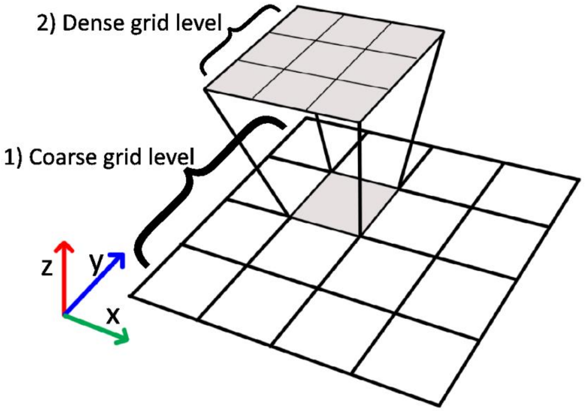
O. Józsa, A. Börcs and <u>*Cs. Benedek*</u>: "Towards 4D Virtual City Reconstruction From Lidar Point Cloud Sequences," ***ISPRS Workshop on 3D Virtual City Modeling***, Regina, Saskatchewan, Canada, May 28-31, 2013, vol. II-3/W1 of ***ISPRS Annals of Photogrammetry, Remote Sensing and the Spatial Information Sciences***, pp. 15-20

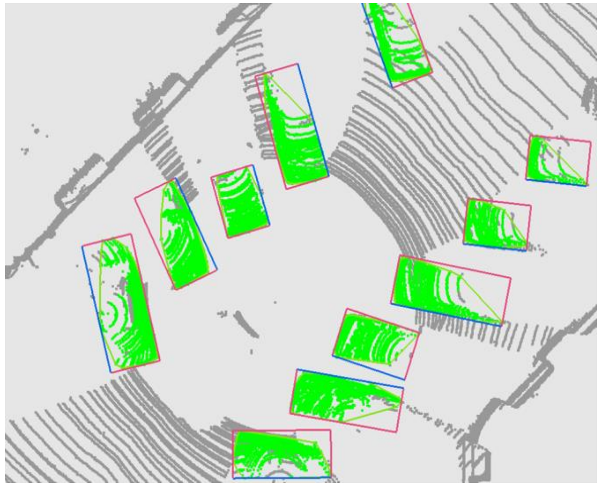
Object separation by a two-level grid model

- Object detection

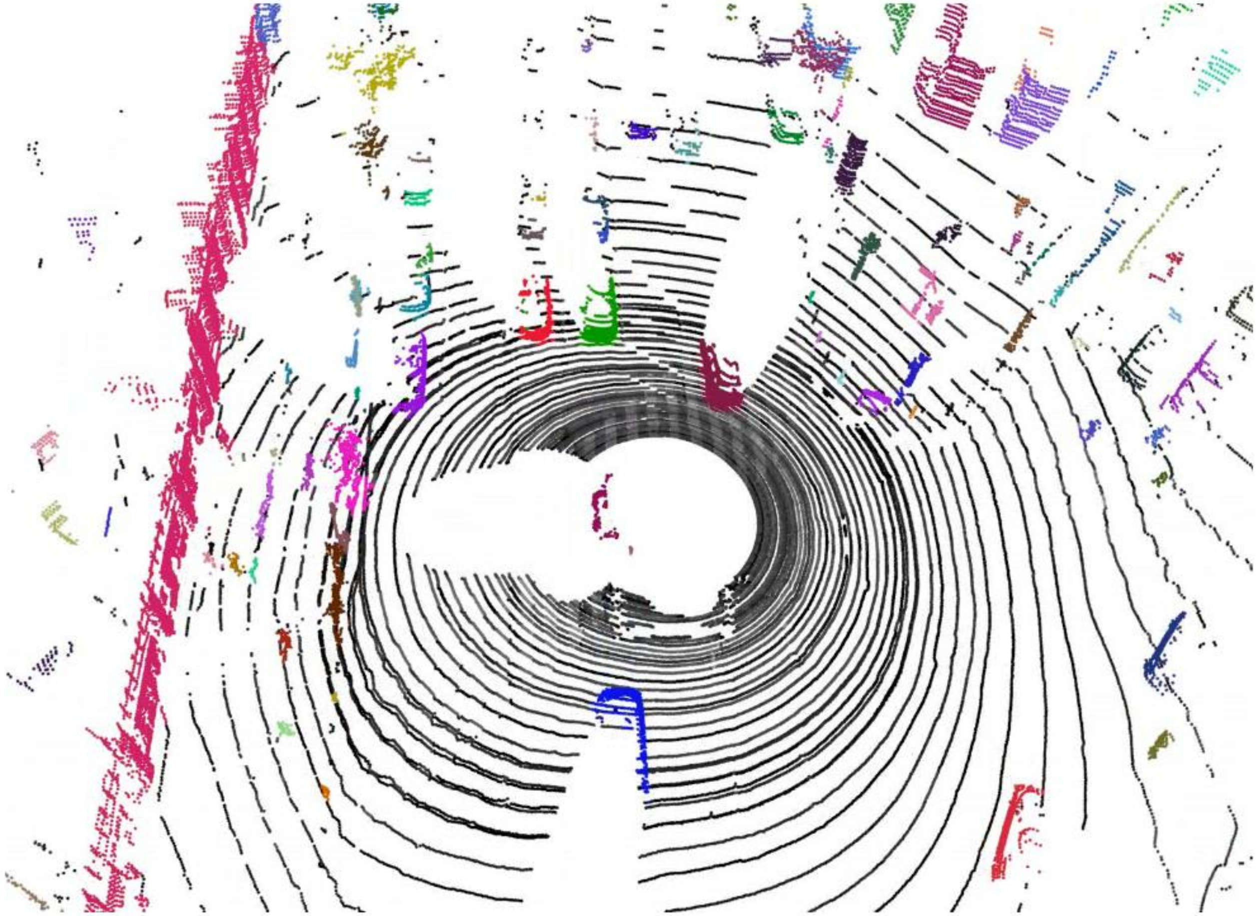
- Separation of close objects with a two-layer hierarchical grid model







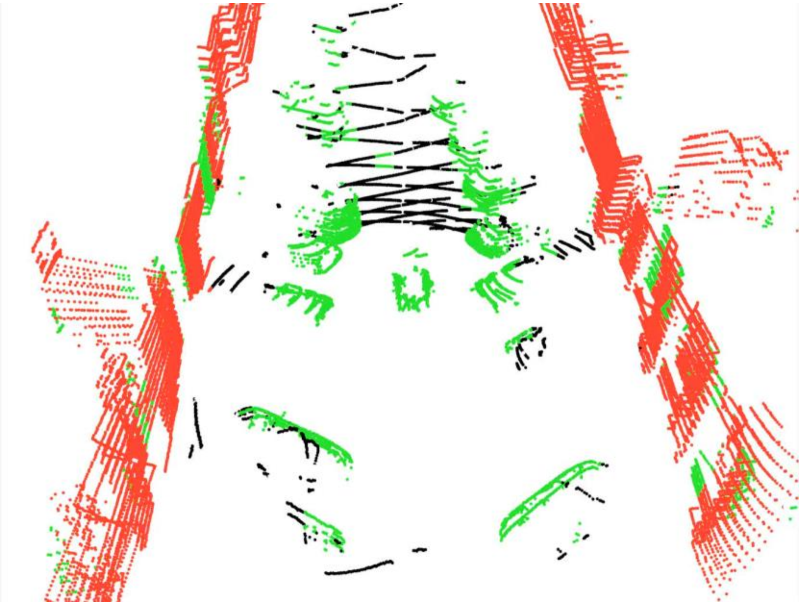
Object detection and tracking with Lidar sensors



A. Börcs, B. Nagy and Cs. Benedek: "Fast 3-D Urban Object Detection on Streaming Point Clouds", ***Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving at ECCV,*** ***Lecture Notes in Computer Science***, Zurich, Switzerland, September 6-12 2014

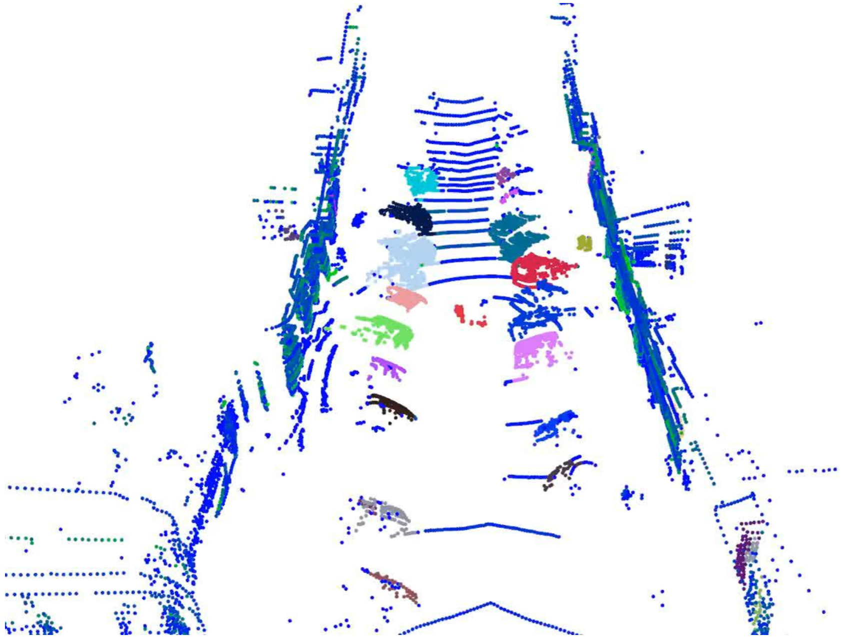
Working with low-resolution Lidars

Point cloud classification



road walls objects

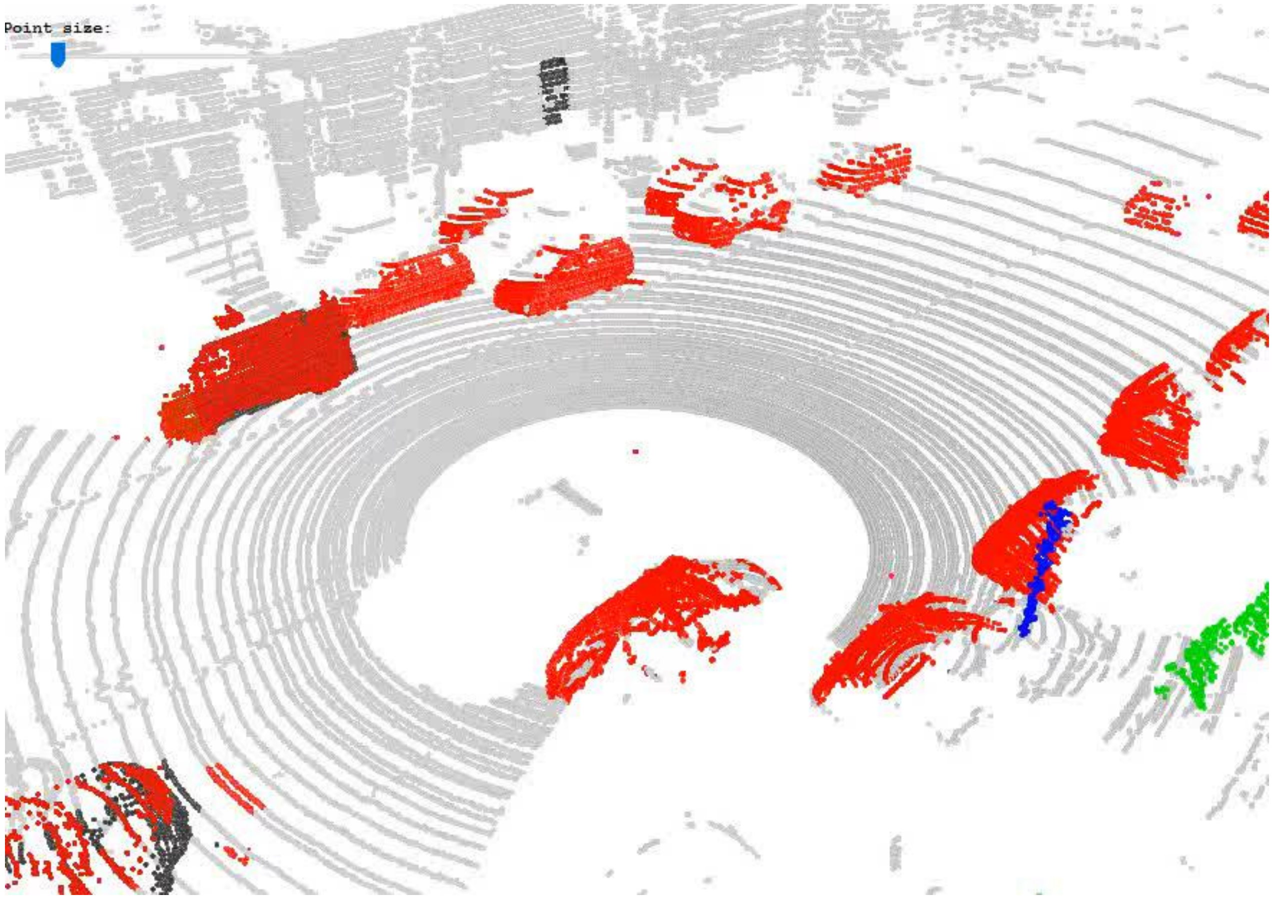
Object tracking in point clouds



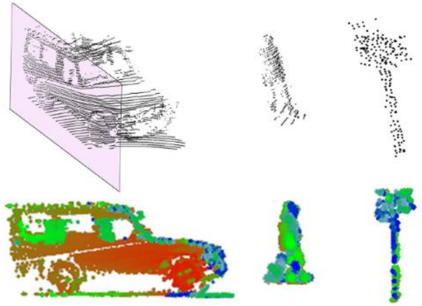
2 x Velodyne VLP-16

A. Börcs, B. Nagy and Cs. Benedek: "Fast 3-D Urban Object Detection on Streaming Point Clouds",
Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving at ECCV,
Lecture Notes in Computer Science, Zurich, Switzerland, September 6-12 2014

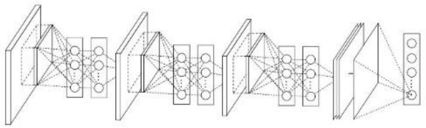
Object classification with Lidar sensors



Depth image representation



Deep neural network

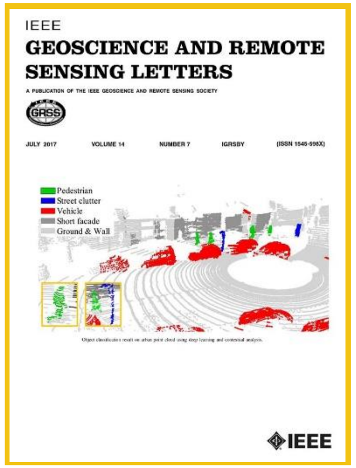


Classification

pedestrian

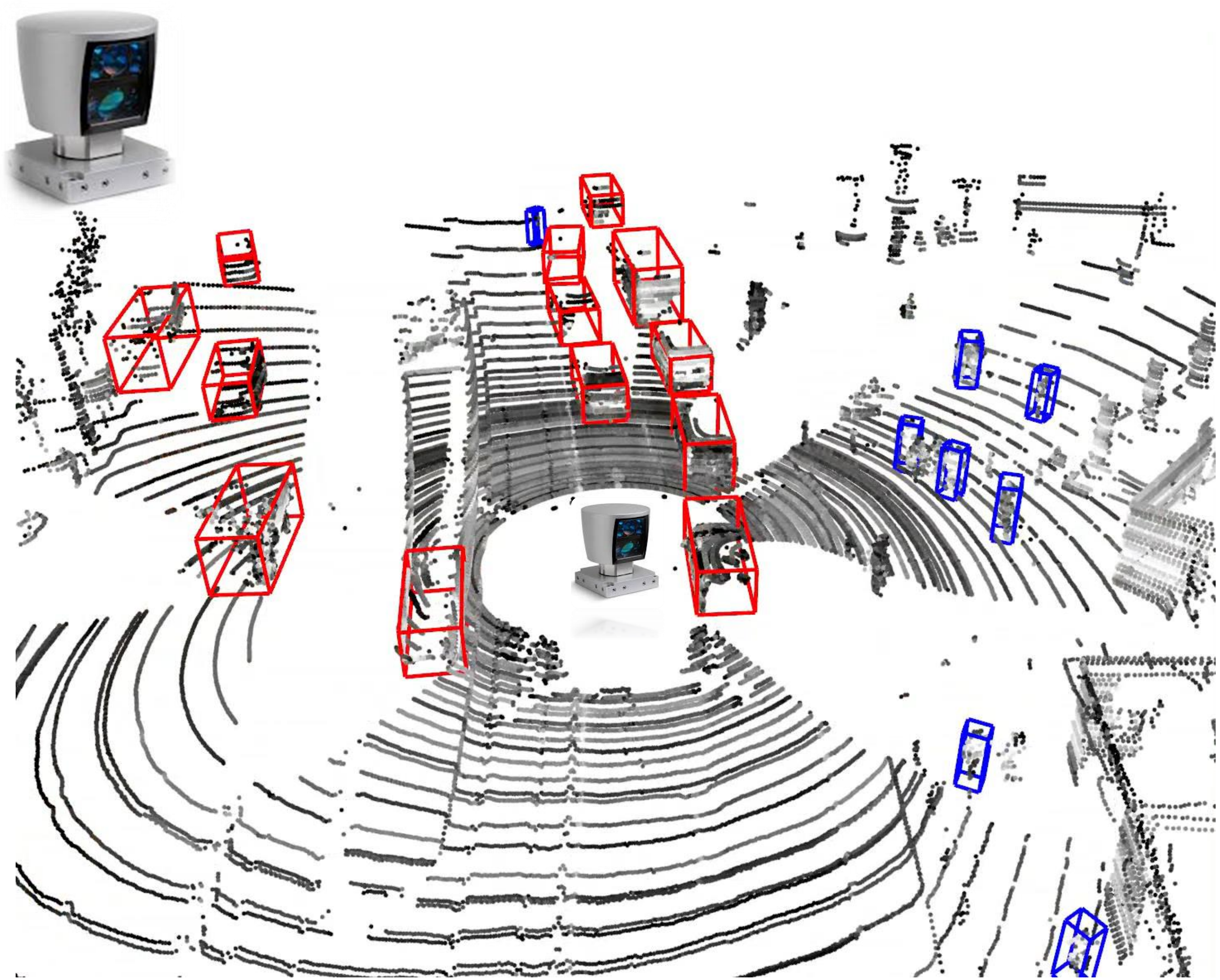
vehicle

street object



A. Börcs, B. Nagy and Cs. Benedek: „Instant Object Detection in Lidar Point Clouds", ***IEEE Geoscience and Remote Sensing Letters***, vol. 14, no. 7, pp. 992 - 996, 2017, IF: 2.892

Lidar based object detection



Lang, A.H. et al.: Pointpillars: Fast encoders for object detection from point clouds. (In: 2019 IEEE Conference on Computer Vision and Pattern Recognition)

Sources of errors:

- Similar appearance of point clouds corresponding to different objects
- Merged object blobs

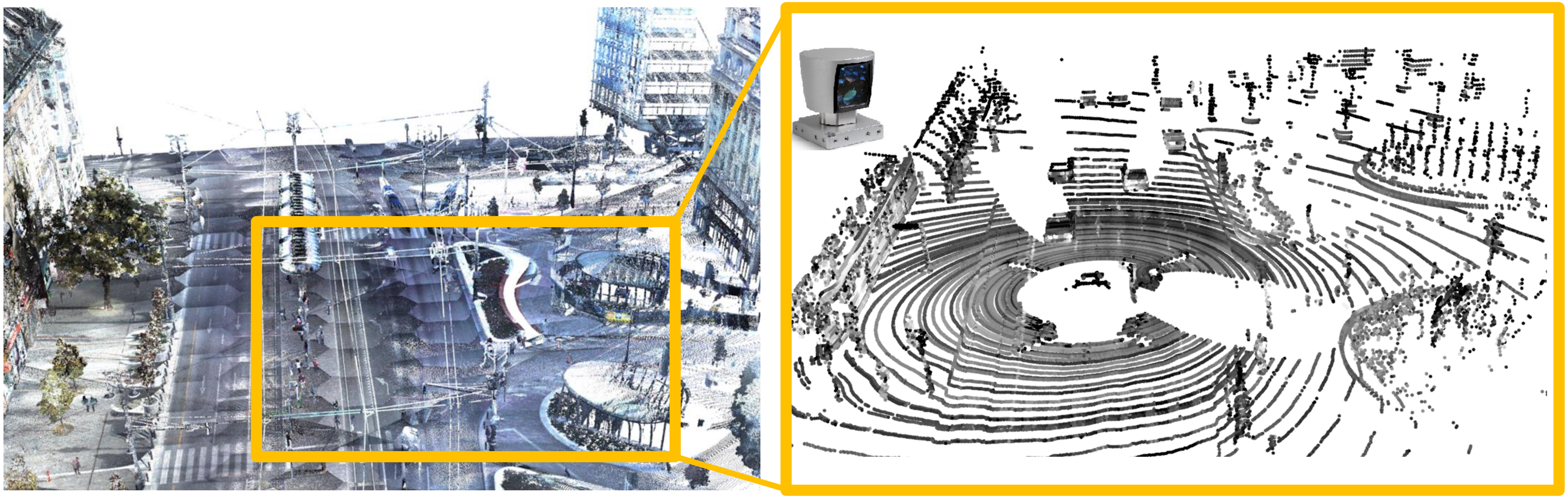
Proposed solution:

- Map based correction

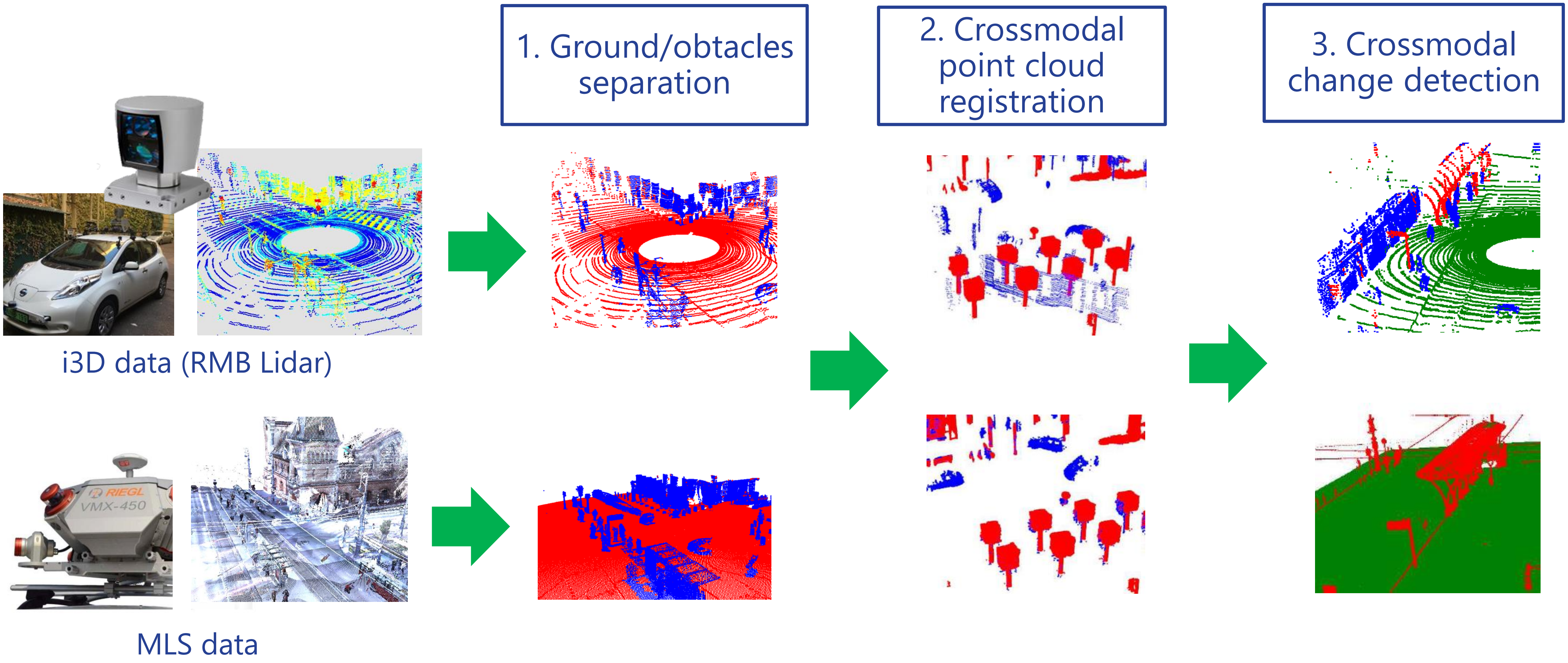
MLS map based vehicle localization and change detection

City Management and Maintenance
Map by Mobile Laser Scanning (MLS)

Intelligent Vehicles (IV):
Instant 3D (i3D) data



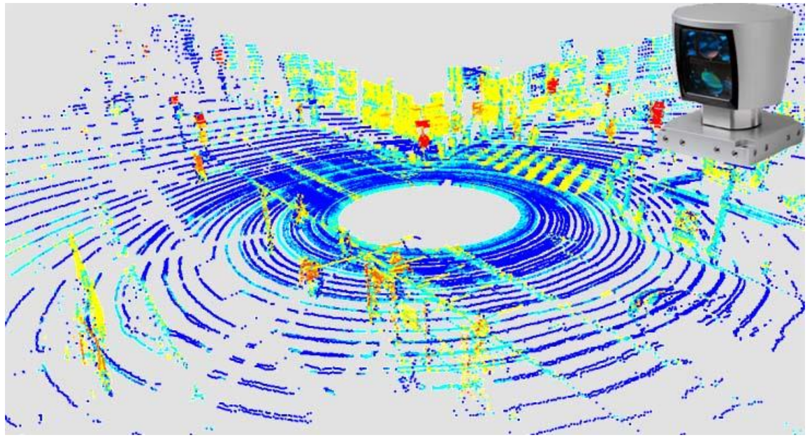
Change detection workflow



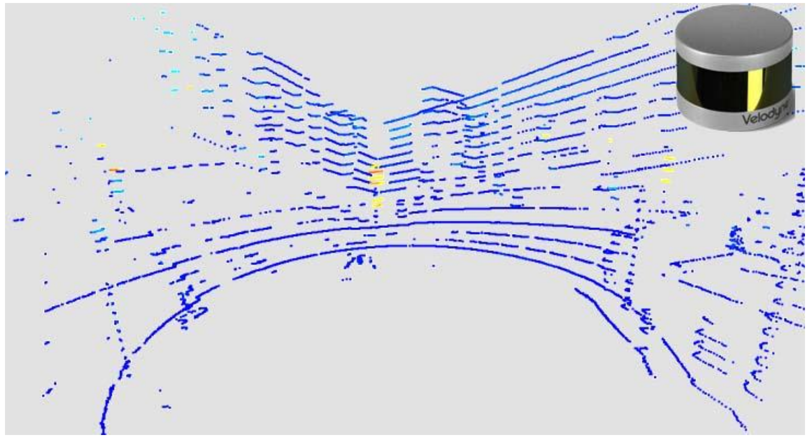
Result of ground-obstacle separation for sensors

Fővám tér, Budapest

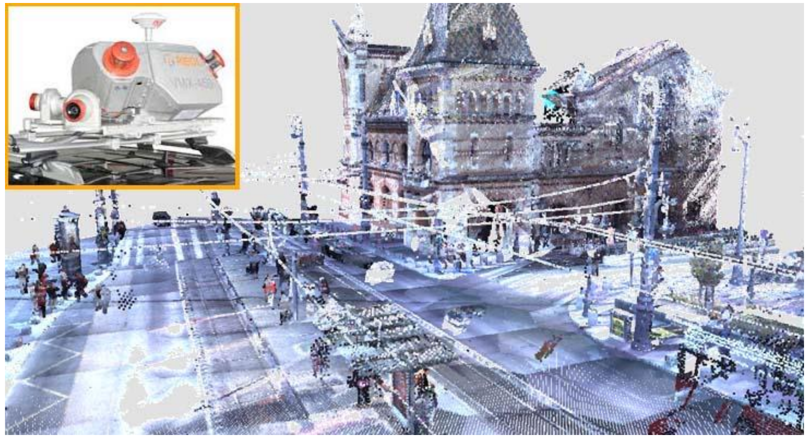
Velodyne HDL64E (i3D)

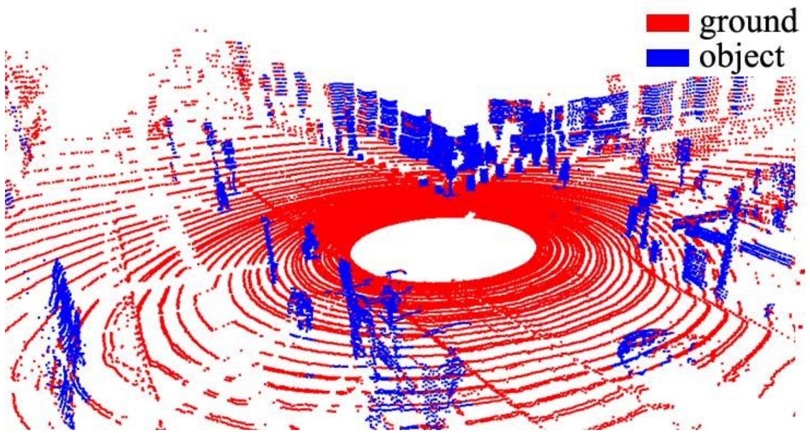


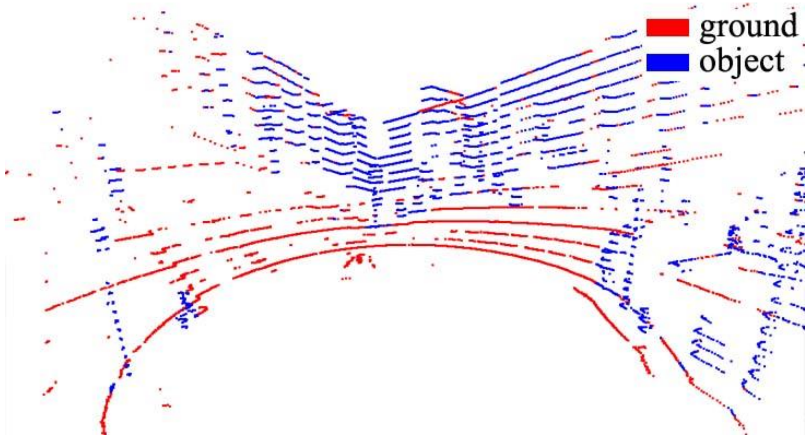
Velodyne VLP16 (i3D)

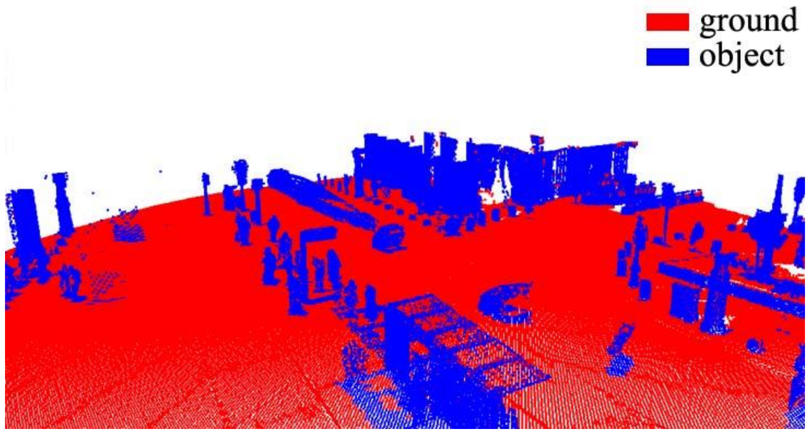


Riegl VMX450 (MLS)







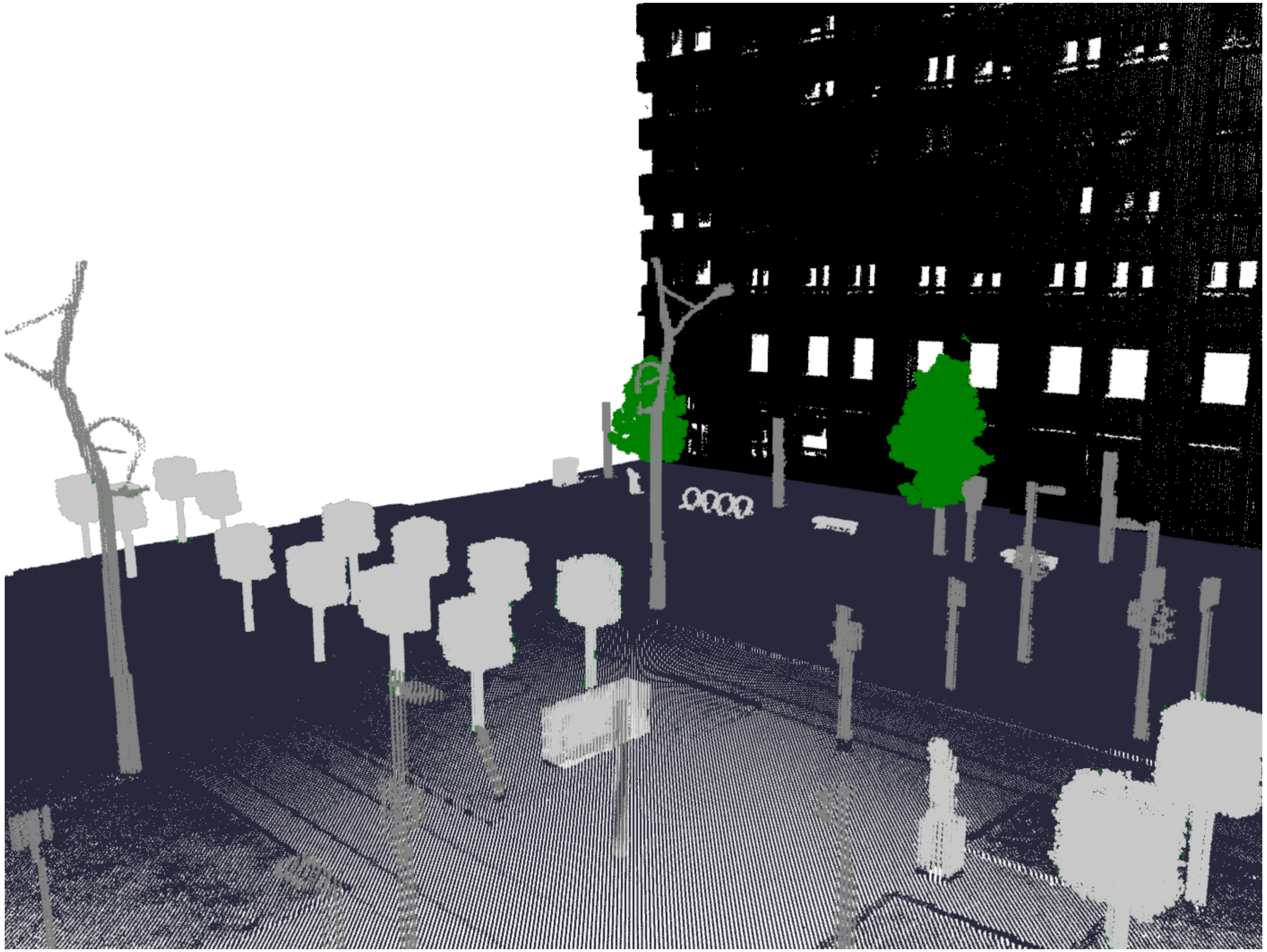


Step 1: Ground-obstacle separation

MLS point cloud based semantic map generation

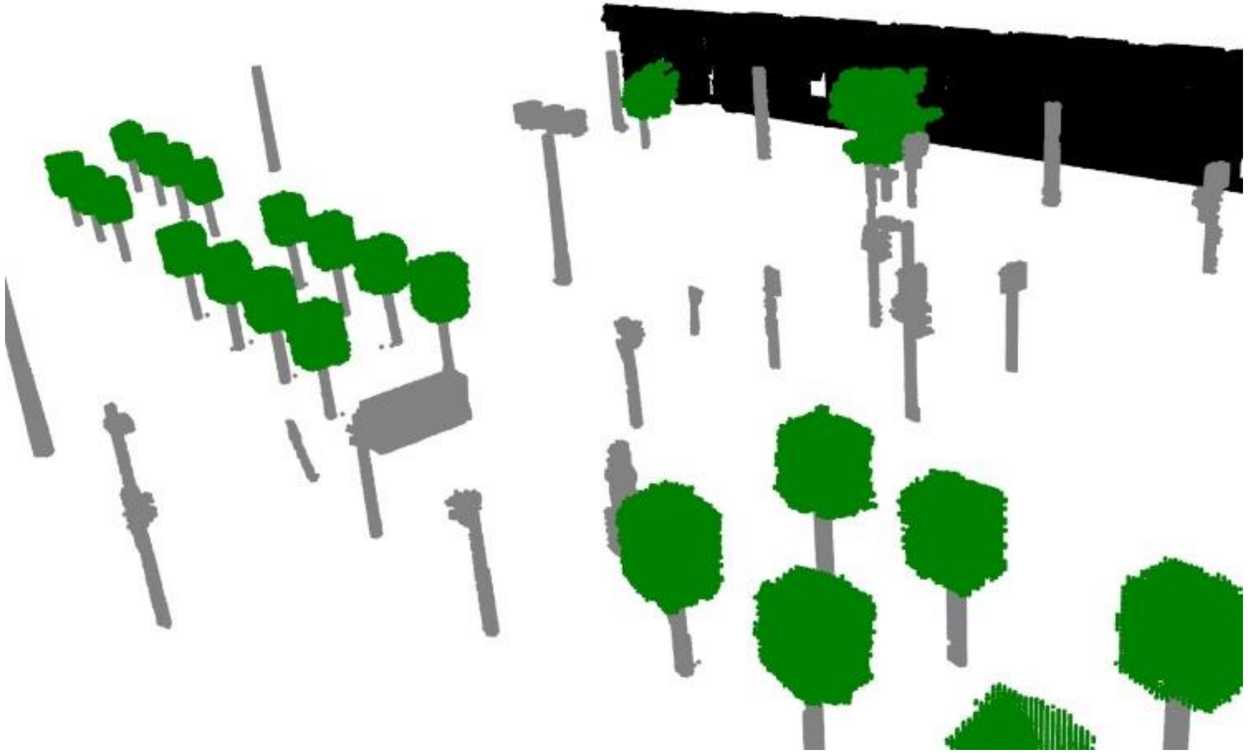
- Point cloud acquired by a Mobile Laser Scanning (MLS) systems
- Handle large point sets (ca. 50 million points)
- Offline data processing
  - Semantic segmentation
  - Filtering
- Reduced point sets (ca. 1 million points)

MLS point cloud based semantic map generation



- Point cloud acquired by a Mobile Laser Scanning (MLS) systems
- Handle large point sets (ca. 50 million points)
- Offline data processing
  - **Semantic segmentation**
  - Filtering
- Reduced point sets (ca. 1 million points)

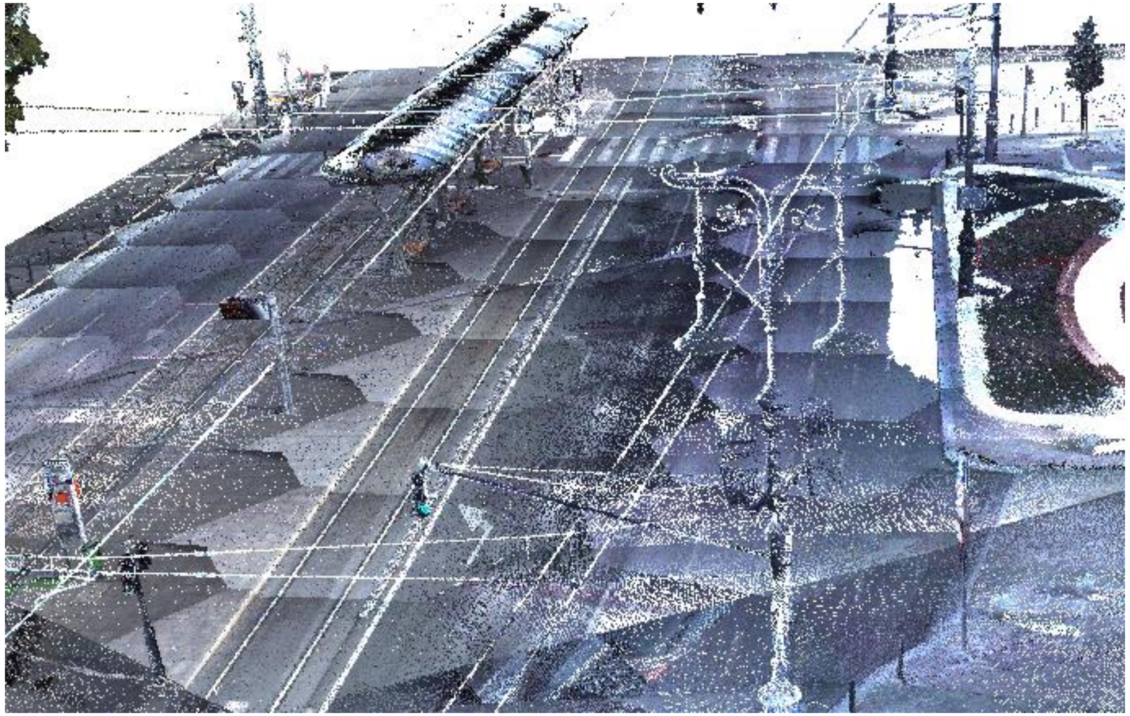
MLS point cloud based semantic map generation

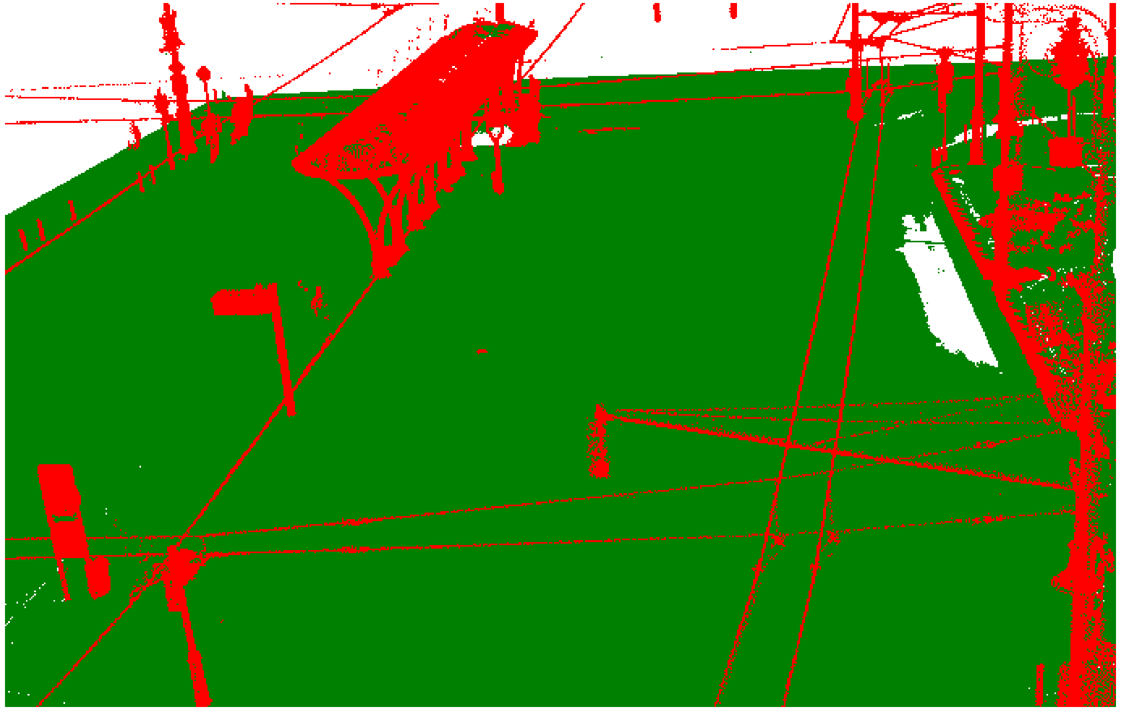


- Point cloud acquired by a Mobile Laser Scanning (MLS) systems
- Handle large point sets (ca. 50 million points)
- Offline data processing
  - Semantic segmentation
  - **Filtering**
- Reduced point sets (ca. 1 million points)

MLS based background point cloud map

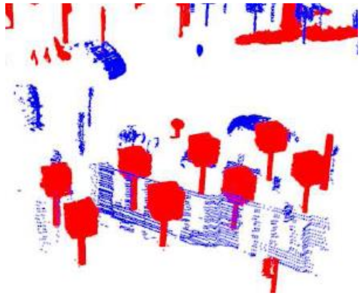
After obstace-ground separation and moving object removal





Step 1: Ground-obstacle separation

Step 2: Proposed point cloud registration workflow



Input: obstacle clouds after ground removal

2.1: Landmark object extraction

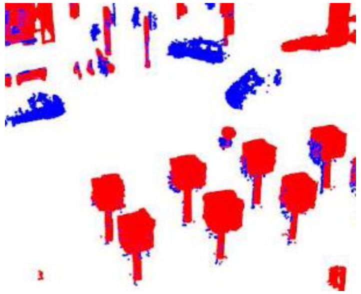
Advanced Euclidean clustering

Hough transform based matching of the extracted object centers

2.2: Object based coarse alignment

2.3: Point level registration refinement

Classical ICP or NDT variant

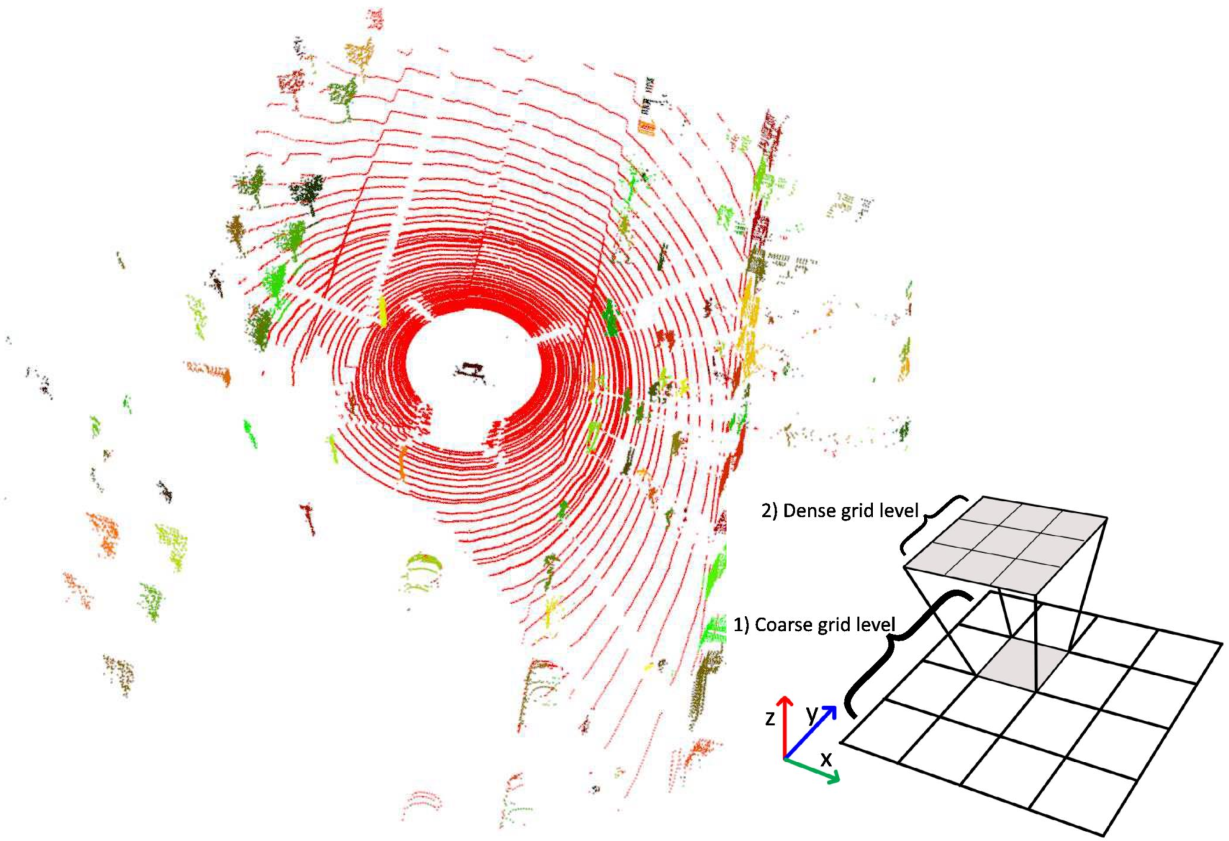


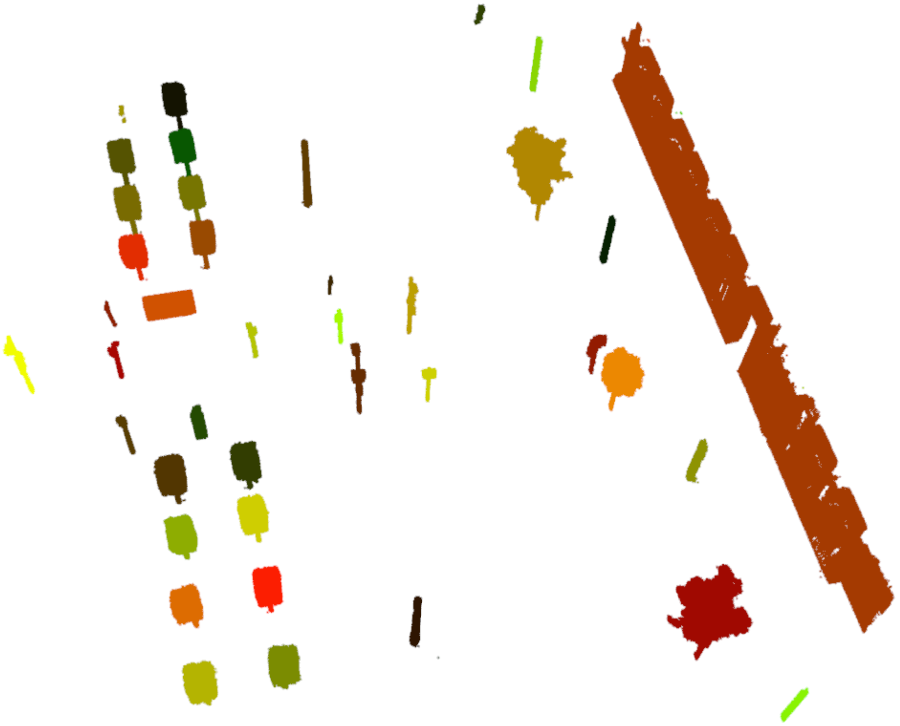
Output: registered point clouds

B. Gálai, B. Nagy and Cs. *Benedek*: "Crossmodal Point Cloud Registration in the Hough Space for Mobile Laser Scanning Data", ***International Conference on Pattern Recognition (ICPR),*** pp. 3363-3368, Cancun, Mexico, 4-8 December 2016

<u>Step 2.1</u>: Landmark object extraction

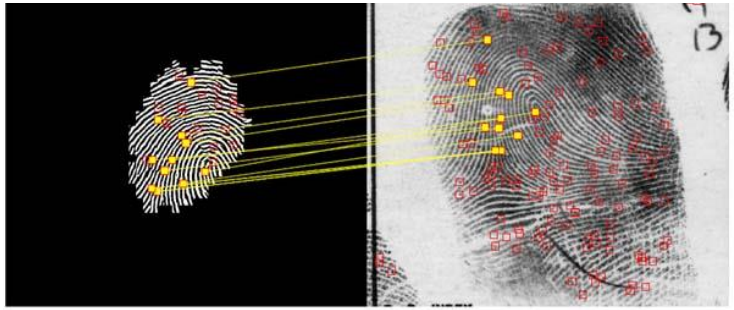
- RMB Lidar point cloud:
  - two layer hierarchical grid model for object separation (see earlier)
  - detected moving objects may be removed (accuracy is not critical)
- MLS point cloud
  - Starting from semantic segmentation map (keeping columns, street furniture elements)
  - Euclidean clustering

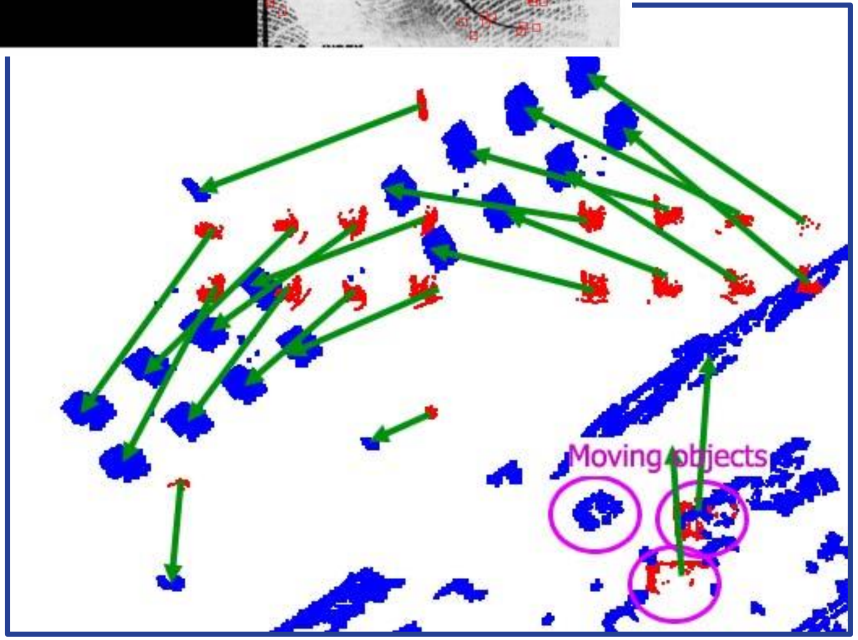




Step 2.2: Object based coarse alignment

- Motivation: fingerprint minutiae matching
  - Generalized Hough transform to estimate the optimal match between the two sets of object centers extracted from the input clouds

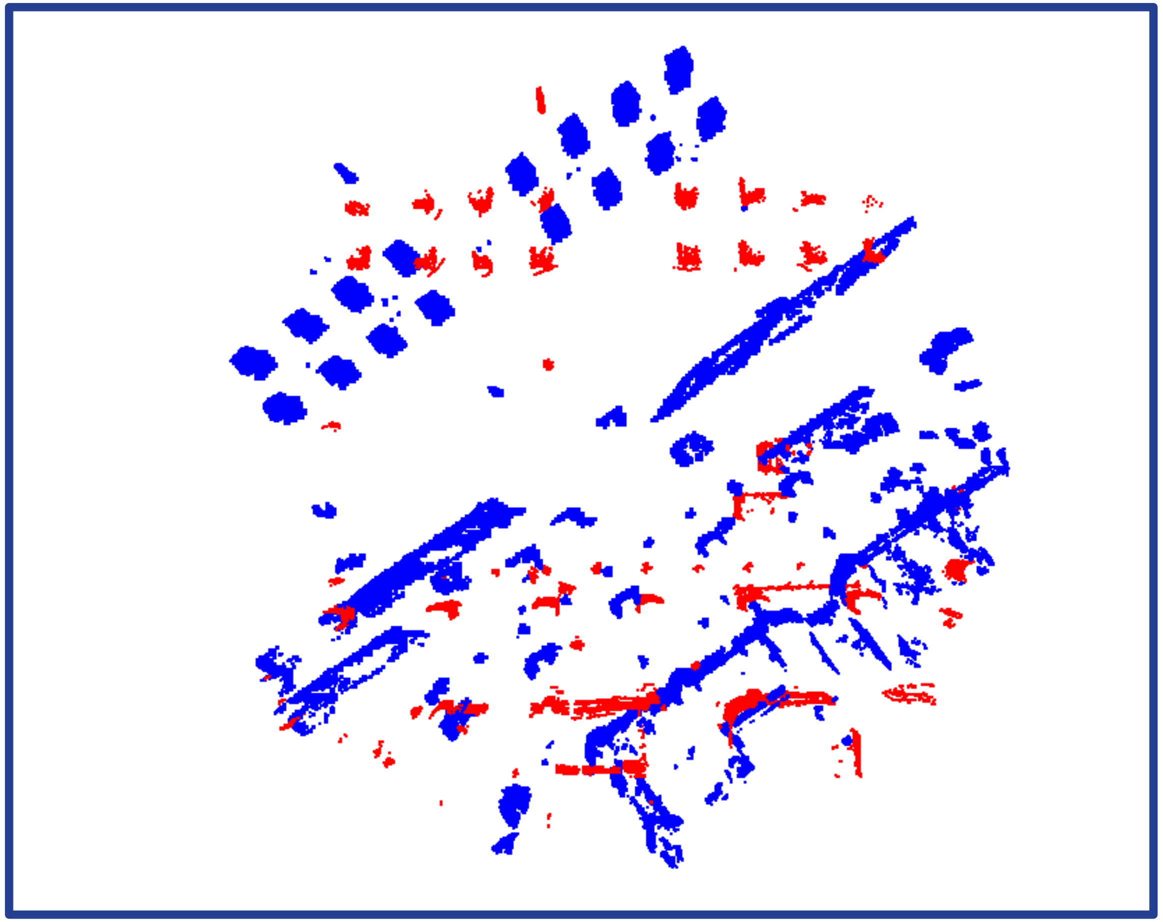


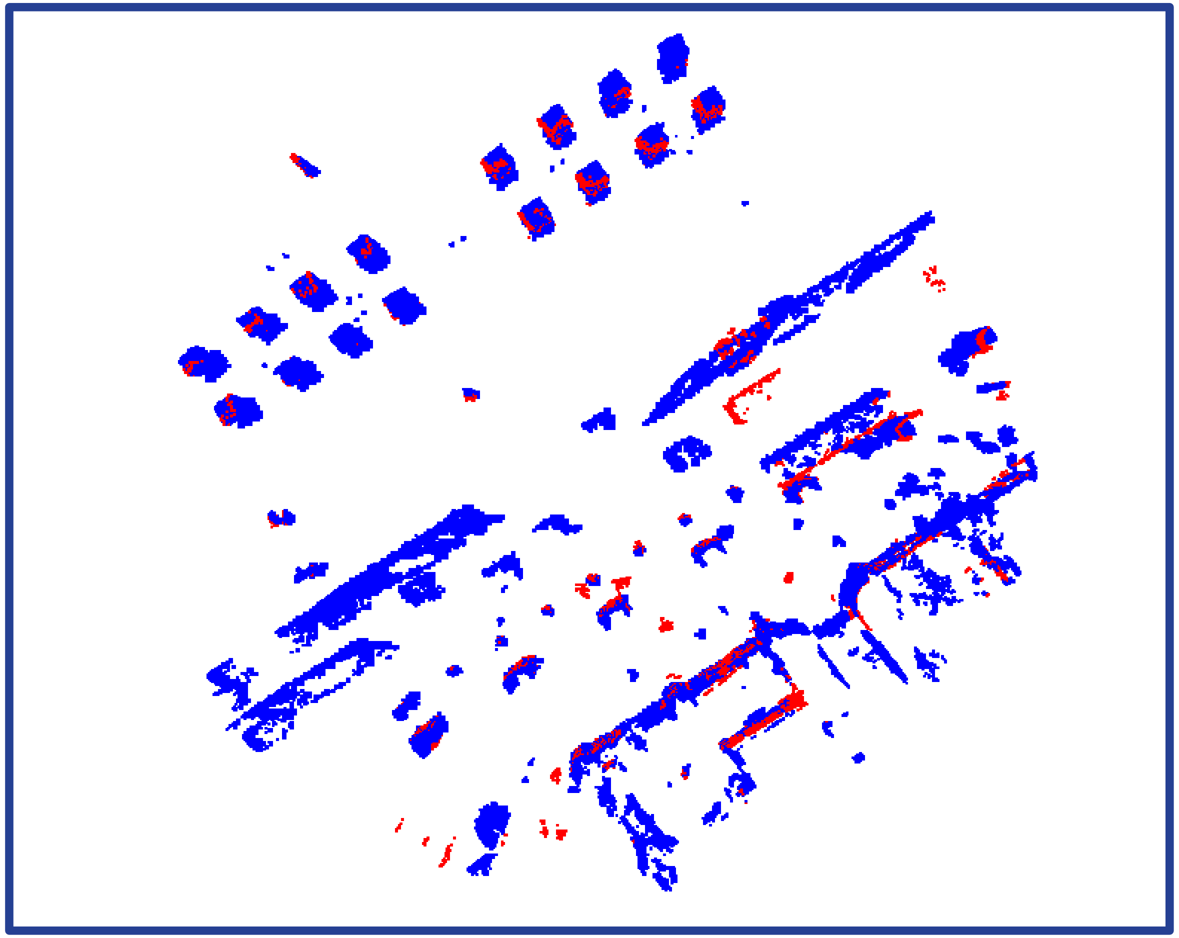



```
procedure Alignment(F1, F2, T)
 C1 <- ObjectDetect(F1)
 C2 <- ObjectDetect(F2)
 Initialize 4D accumulator array A
 for all o1 in C1 do
  for all o2 in C2 do
   for a in [0, 359] do
    o1’ <- Rot(a) * o1
    (dx, dy, dz) <- o2 – o1’
    A[dx,dy,dz,a] <- A[dx,dy,dz,a] + 1
   end for
  end for
 end for
 a, dx, dy, dz <- FindMax(A)
 F1, T <- TransformCloud(F1, a, dx, dy, dz)
end procedure
```

Coarse alignment result

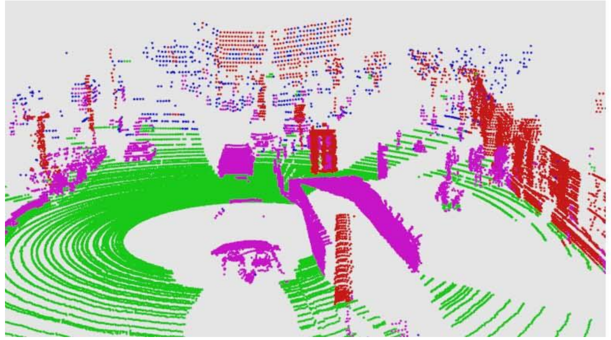
- Algorithm handles large initial rotation error
- Point clouds to be matched may have significantly different resolution and density characteristics
- Point clouds may be recorded at different times: several non-consistent object samples

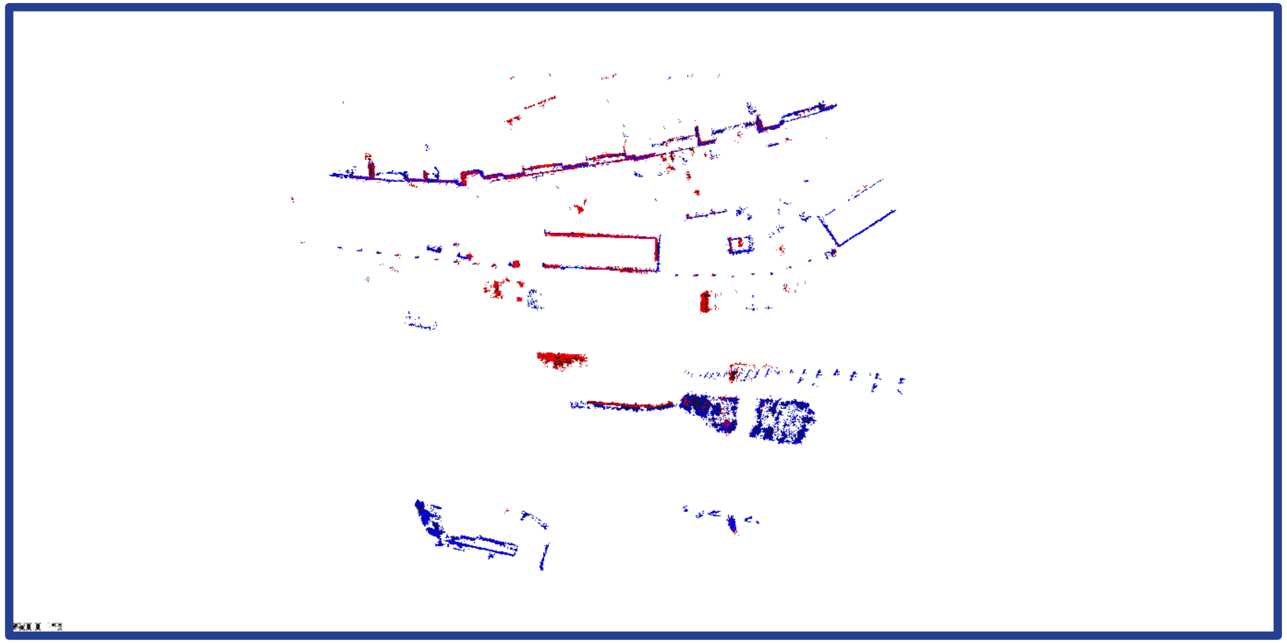
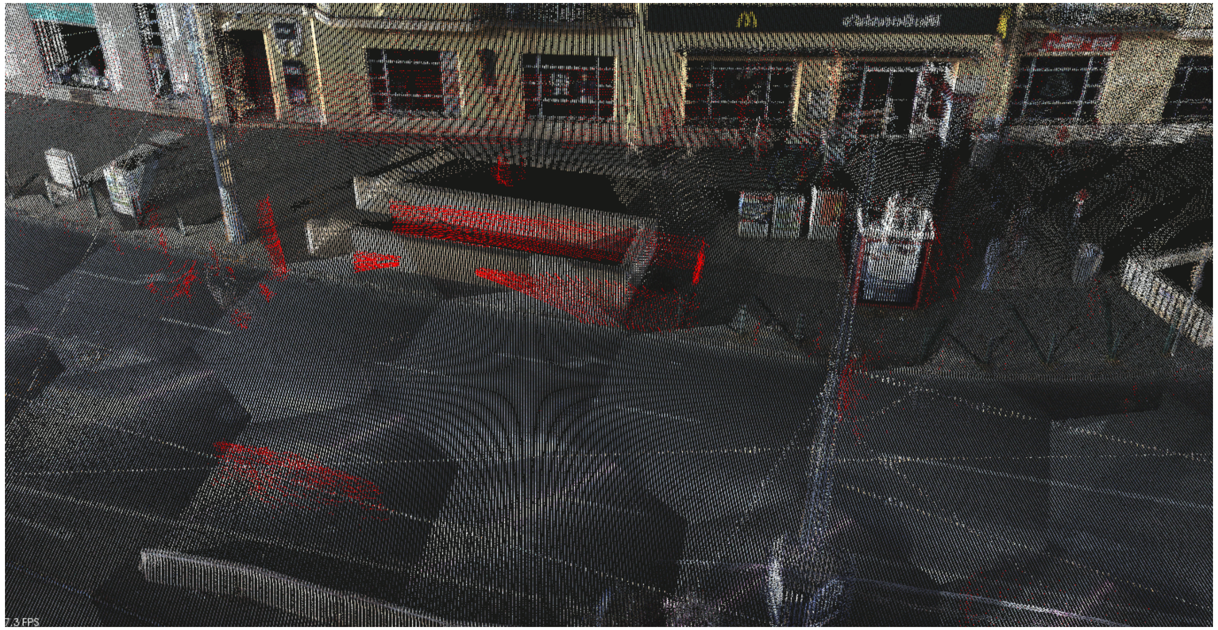




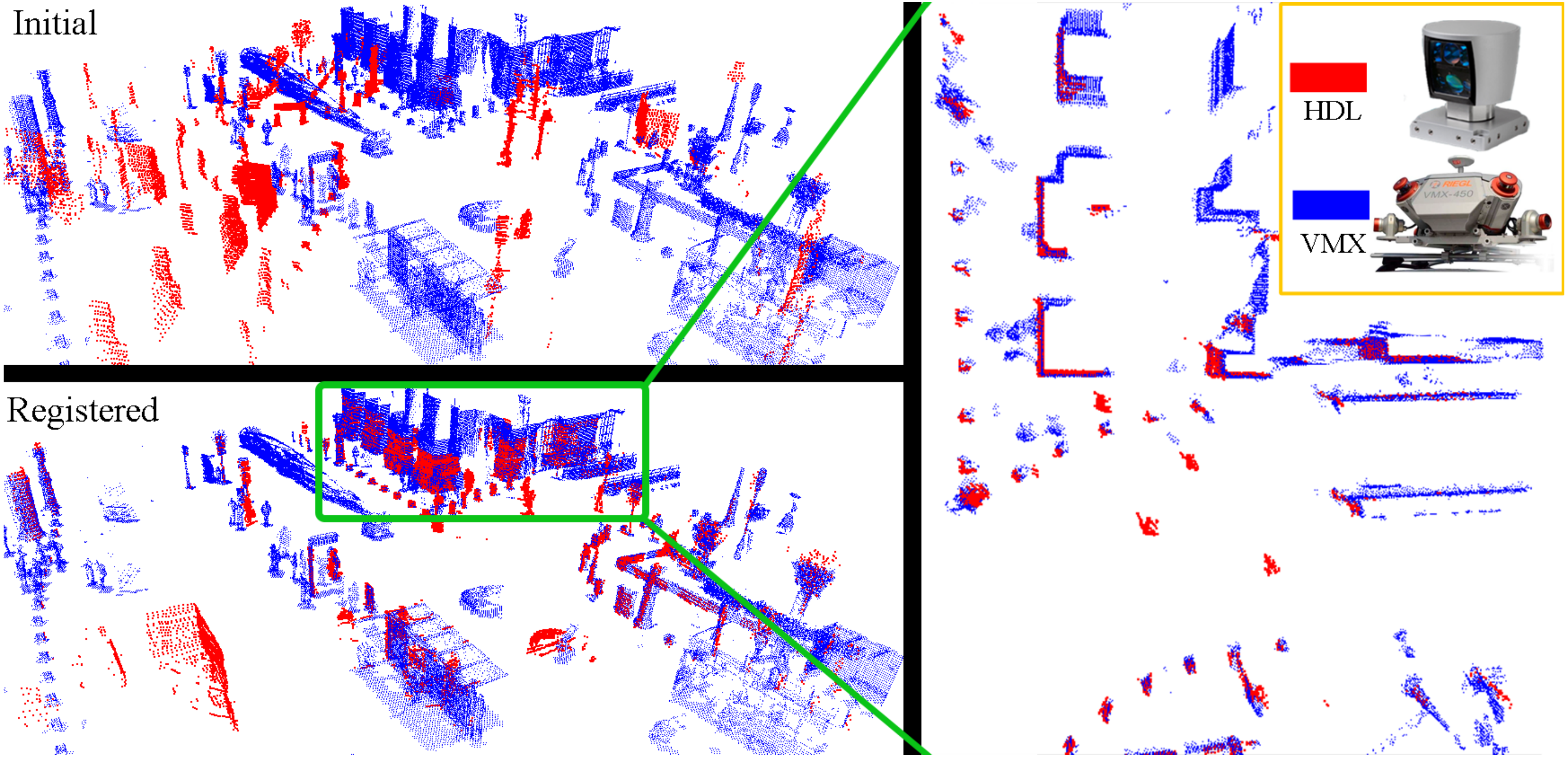
<u>Step 2.3:</u> Point level registration refinement

- *Normal Distributions Transform (NDT)* or *Iterative Closest Point (ICP)* applied on the obstacle cloud

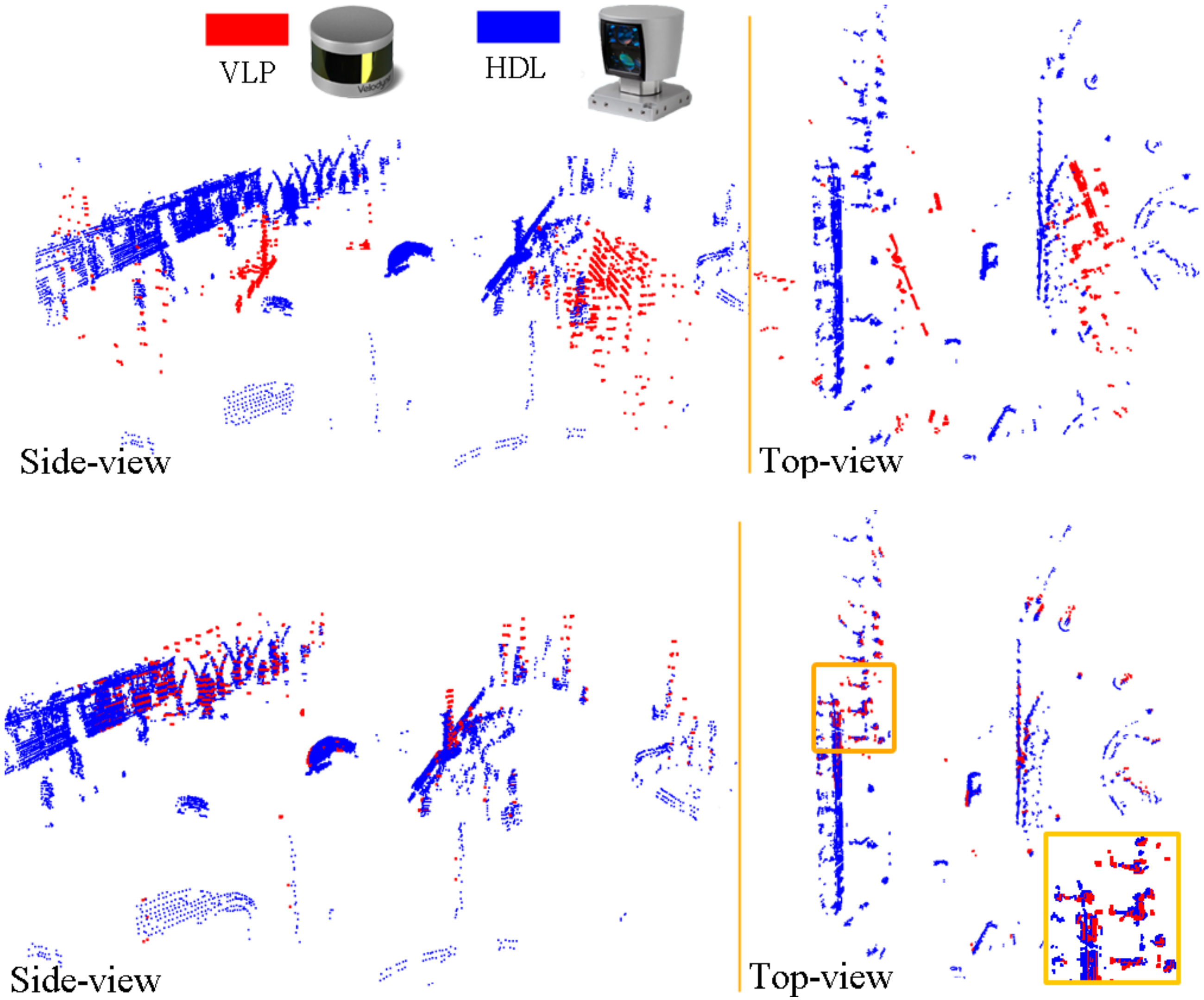


Registration results for high definition inputs



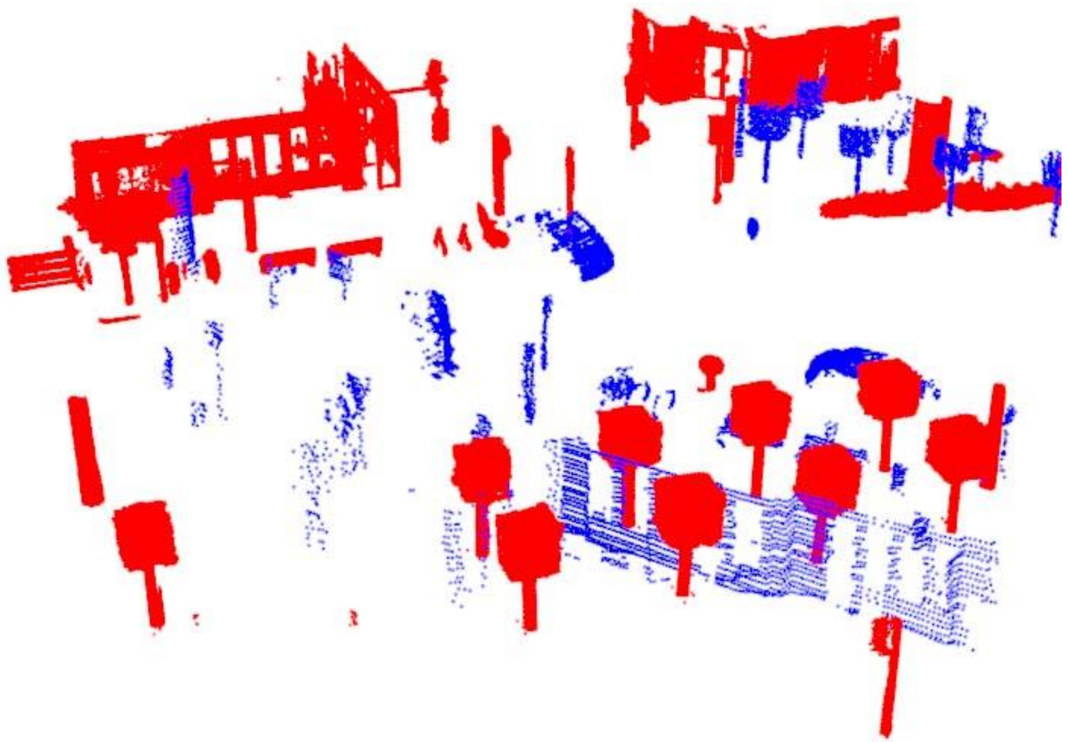
Registration results between VLP16 and HDL64 inputs

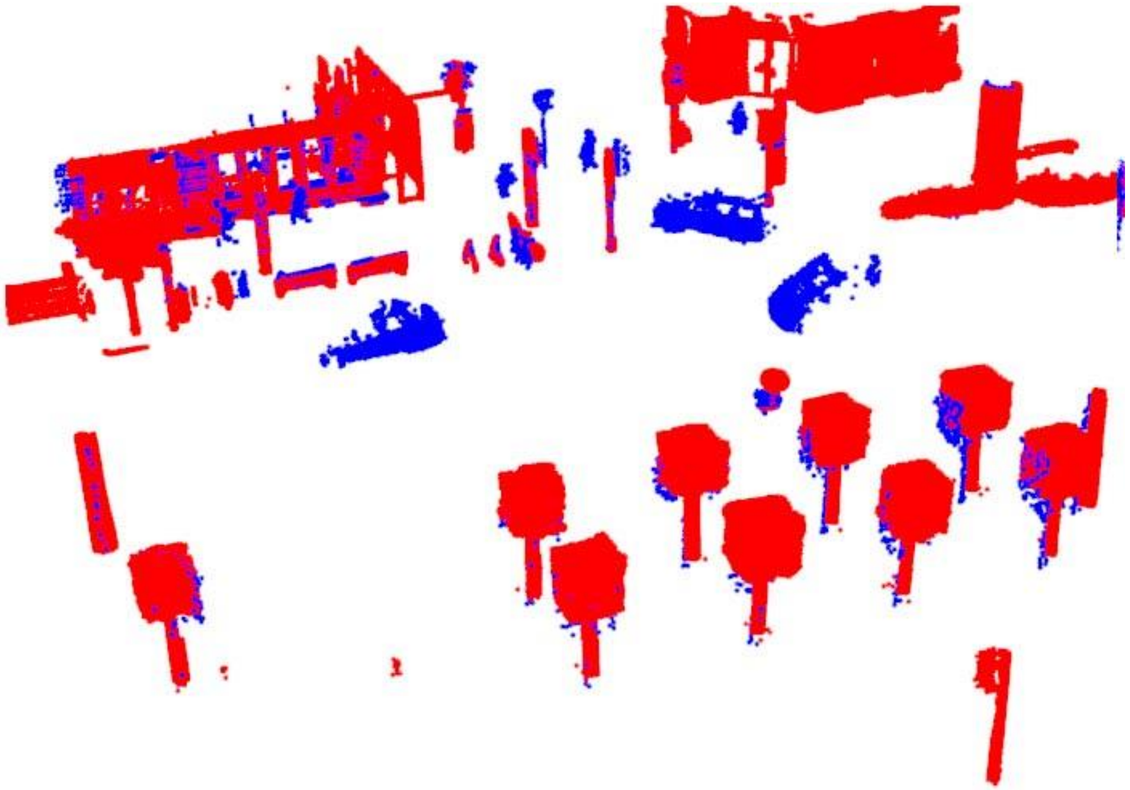


- Algorithm can be also used to auto-calibrate different senors on the same platform
- For even sparser point clouds (e.g. 8 or 4 beam Lidars) support of IMU is recommended

Point cloud registration results

Deák tér, Budapest



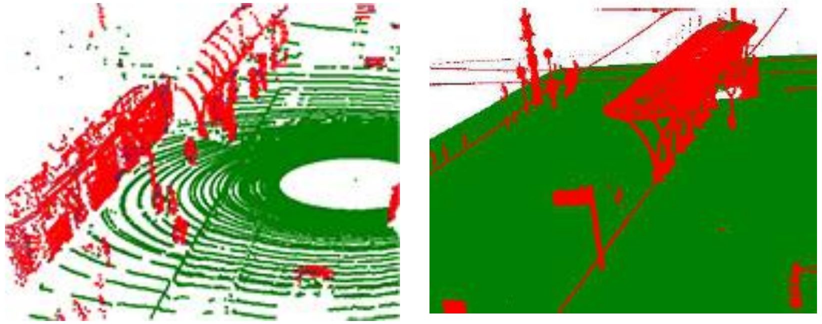


Step 2: Crossmodal point cloud registration

Step 3: Crossmodal change detection

i3D vs. MLS point cloud matching

Change detection workflow:



Input: registered i3D and MLS obstacle clouds

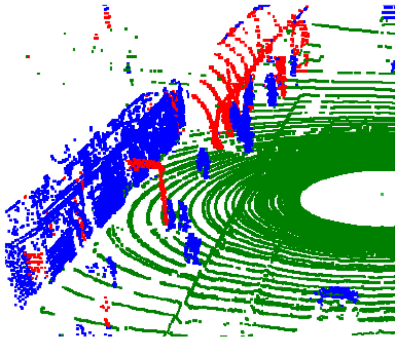
3.1: Range image creation

3.2: MRF based change detection in the range image domain

3.3: Label backprojec-tion to point cloud

Dynamic

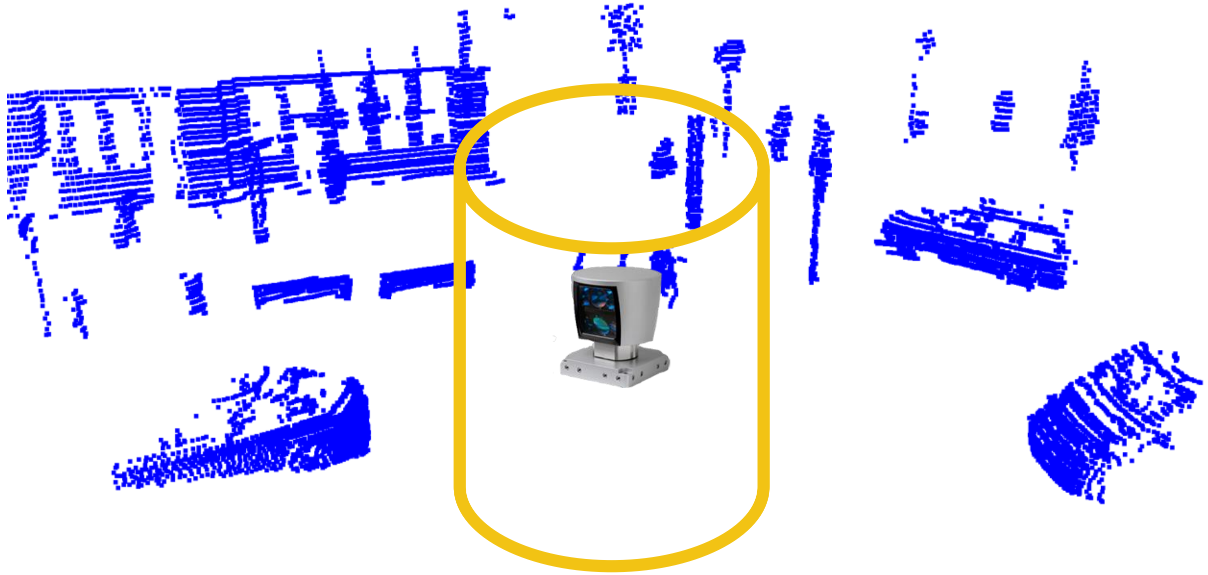
Static



Output: separtion of dynamic and static regions of the i3D obstacle cloud

Step 3.1 Range image formation

- Velodyne range image:
  - Projection on a cylinder around the main axis of the RMB Lidar sensor
  - Interpolation of missing data of the range image due to
    - inhomogeneous point clouds
    - quantization errors of the rotation angle



- MLS range image:
  - Generated from the 3D MLS point cloud with ray tracing
    - available global position and orientation of the RMB Lidar after registration
    - simulated rays emitted into the MLS cloud from the moving platform’s center position with the same vertical and horizontal resolution as the RMB Lidar scanner

3.1 Range image formation

Raw Velodyne (i3D) range image

Filtered & interpolated Velodyne (i3D) range image

MLS range image

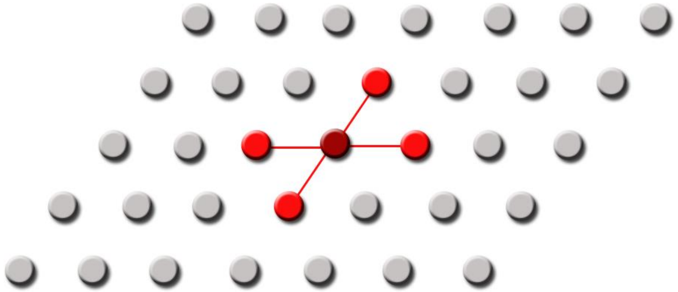
3.2 Change detection in the range image domain

Filtered & interpolated Velodyne (i3D) range image

MLS based range image from the Velodyne's position

MRF based change mask in the range image domain

- Markov Random Field (MRF)



  - Energy minimization (potts modell)

$$E = \sum_{s\in S} V_D\left(\delta_s^{i3D} | l_s\right) + \sum_{s\in S}\sum_{r\in N_s} \beta \cdot 1\{l_s \neq l_r\}$$

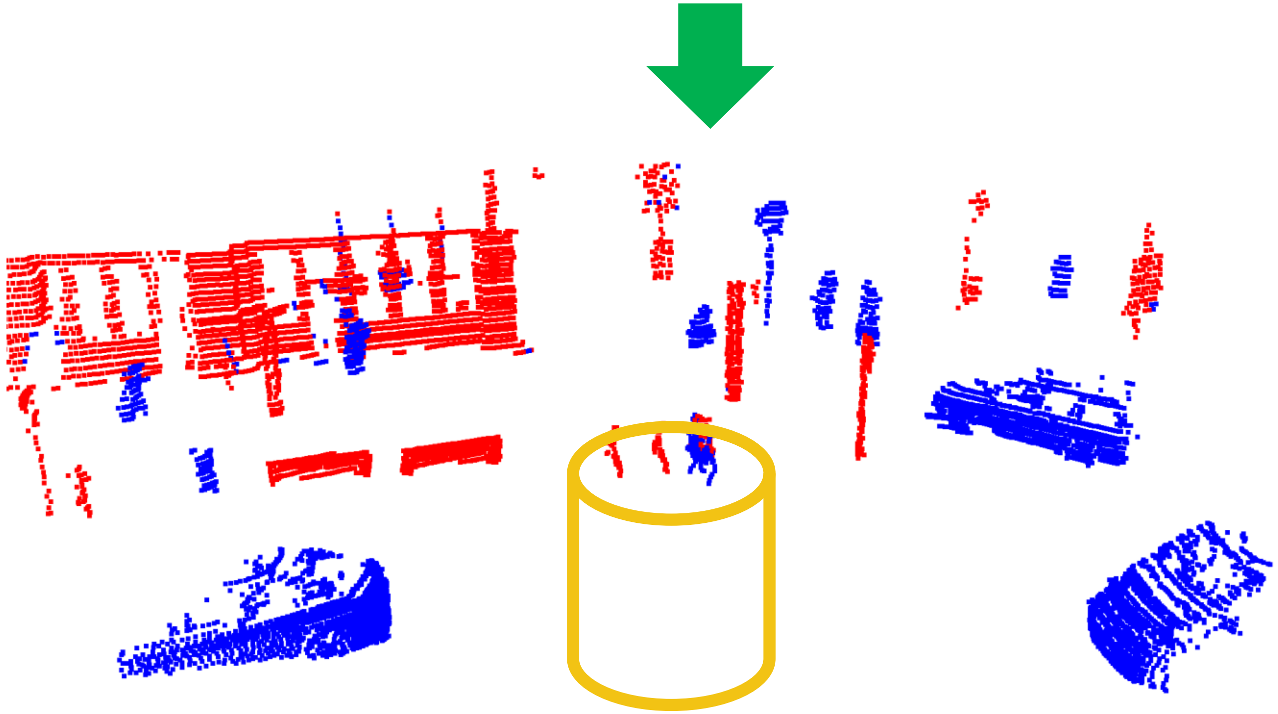
__Data term:__ consistent range values in the input images

__Prior term:__ smooth change mask

3.2 Label backprojection

Filtered & interpolated Velodyne (i3D) range image

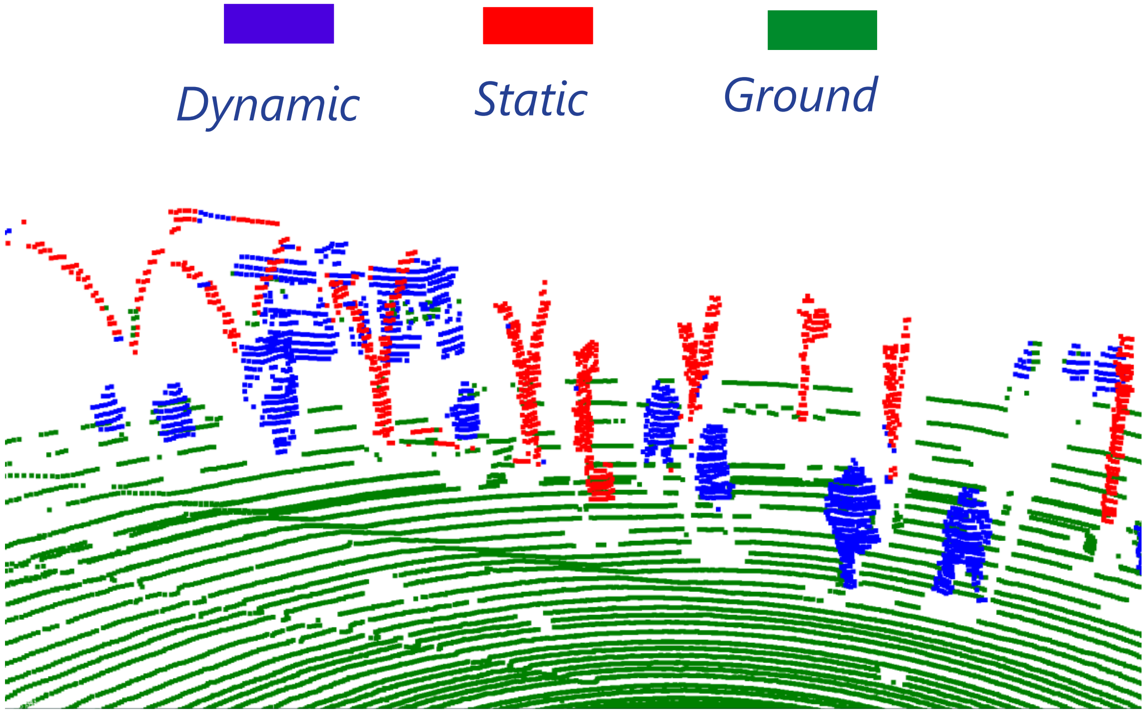
MRF based change mask in the range image domain



Back projection of the 2D change mask to the 3D Velodyne (i3D) point cloud

Change detection resulst in a tram stop

Reference background MLS pont cloud

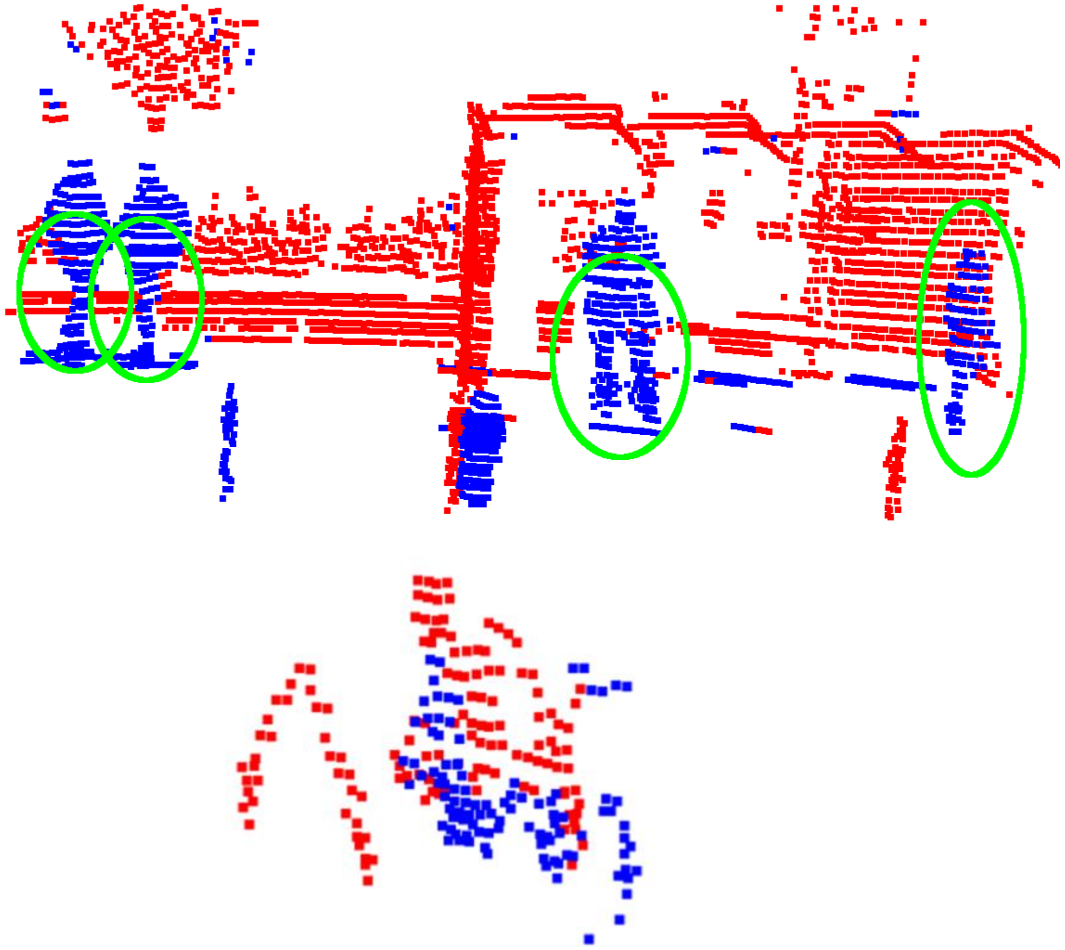


Classified instant 3D Lidar point cloud

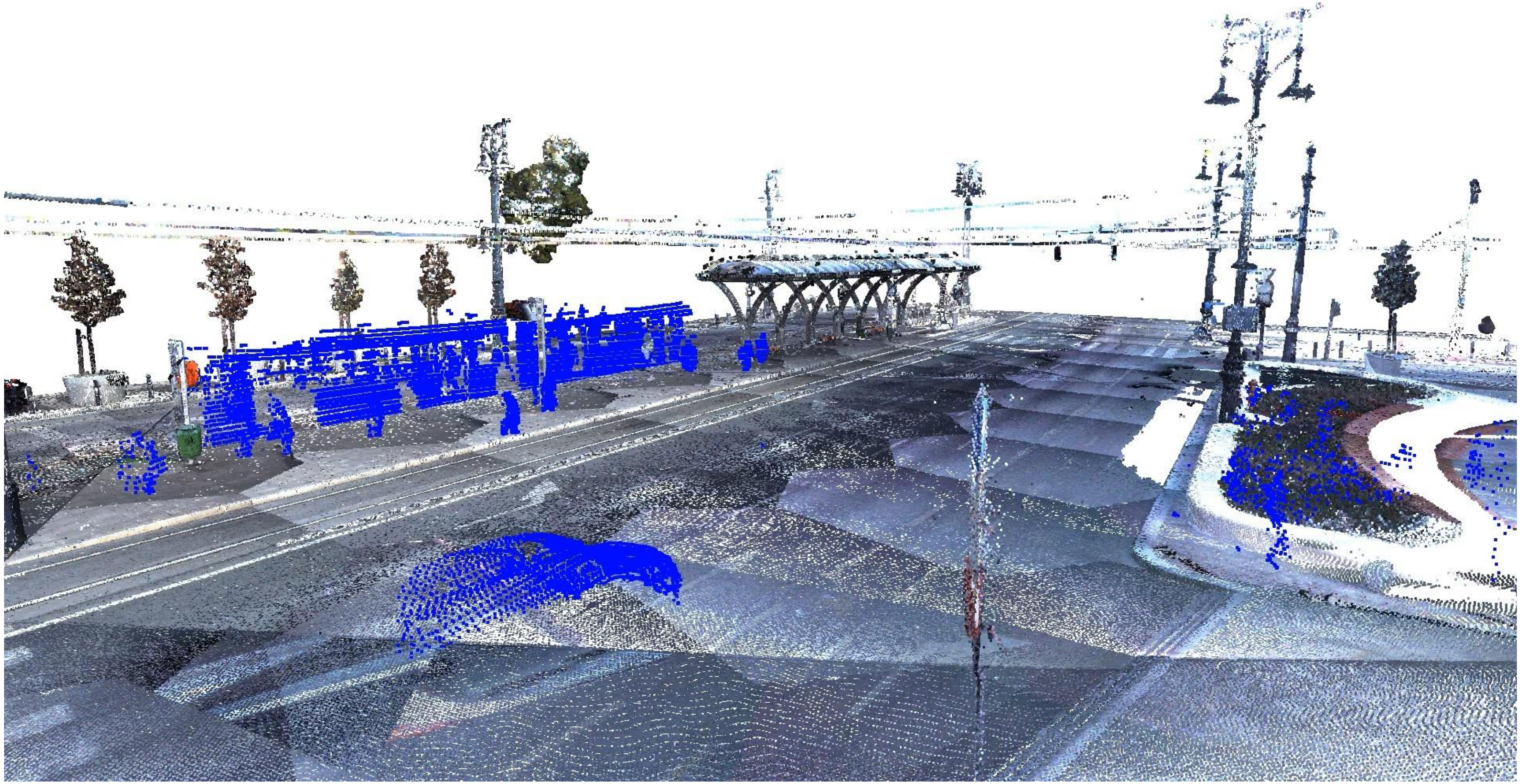
Comparison to a voxel based technique in a sidewalk area

Voxel based (VOX)

Proposed MRF-range image based

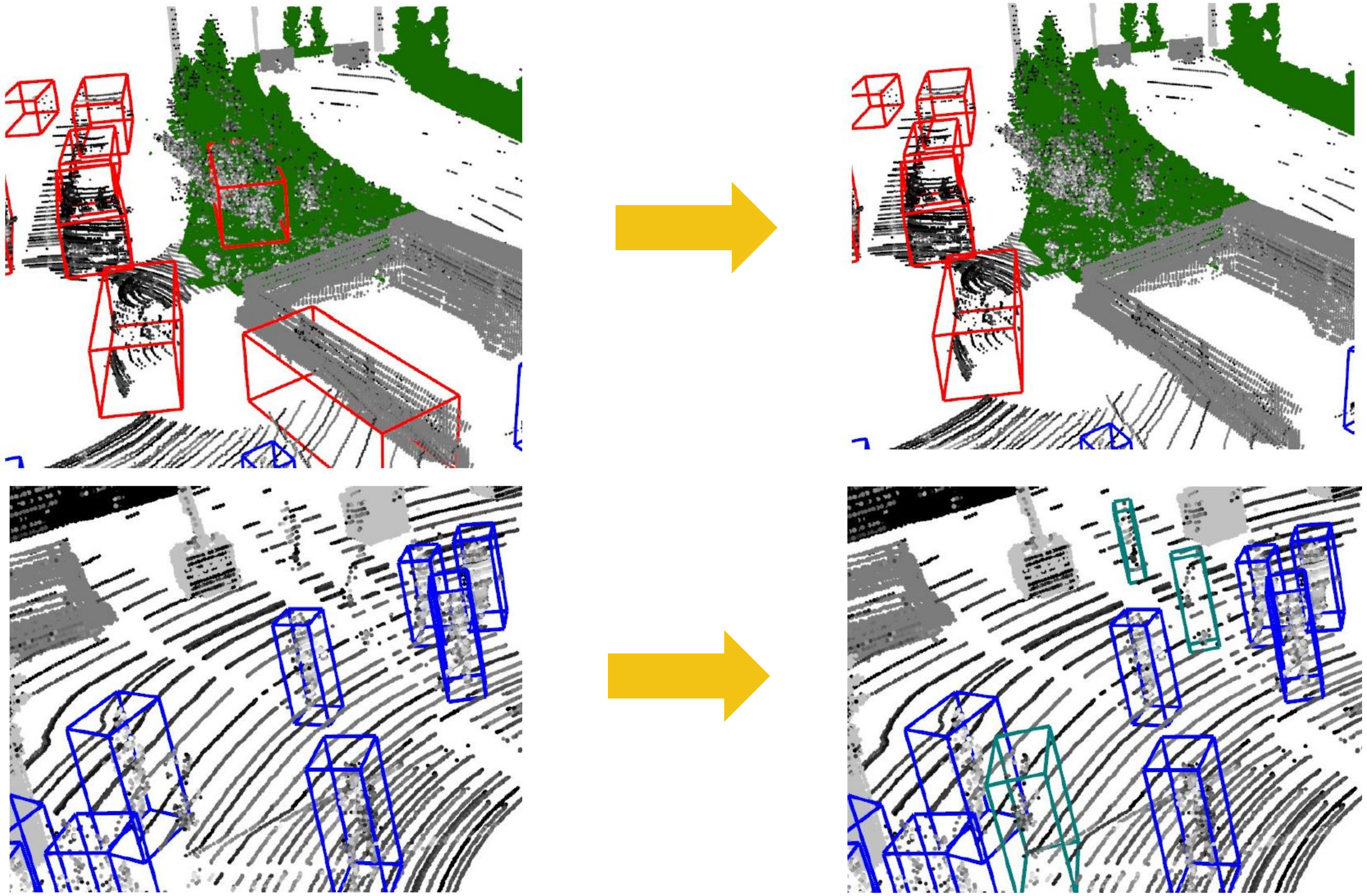


Velodyne - MLS data registration
Synthetized view for geo-referred moving object detection



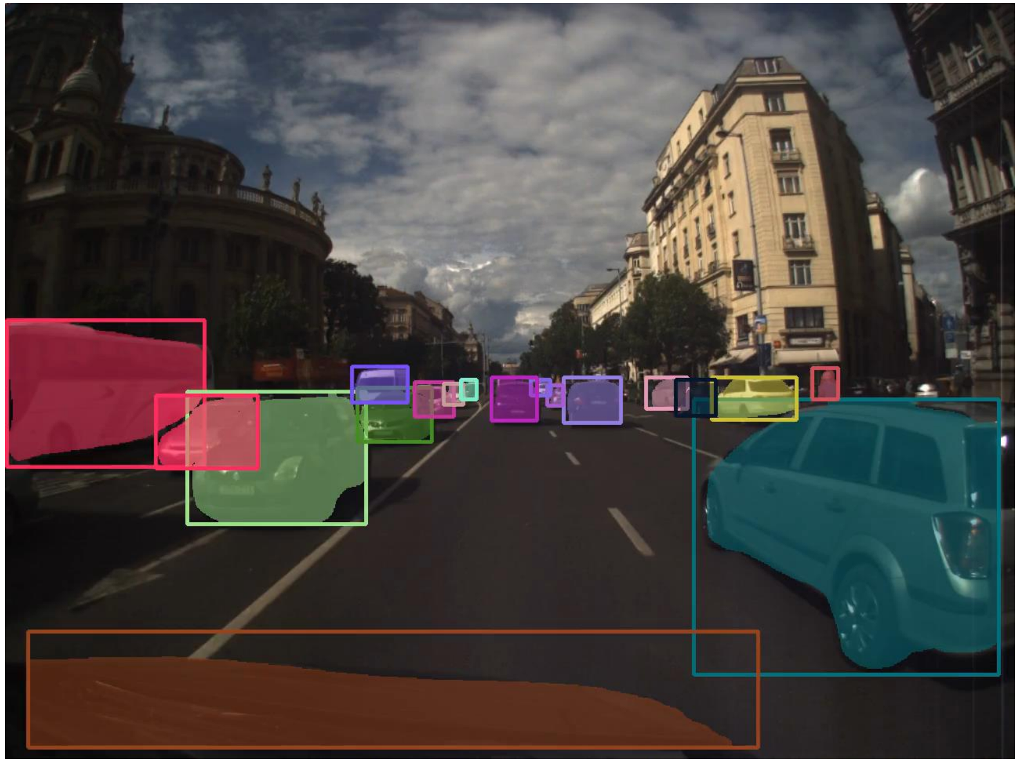
B. Gálai, and *Cs. Benedek*: "Change Detection in Urban Streets by a Real Time Lidar Scanner and MLS Reference data", ***International Conference on Image Analysis and Recognition (ICIAR)***, Montreal, Canada, July 5-7, 2017, vol. 10317 of *Lecture Notes in Computer Science*, pp. 210-220, Springer, 2017

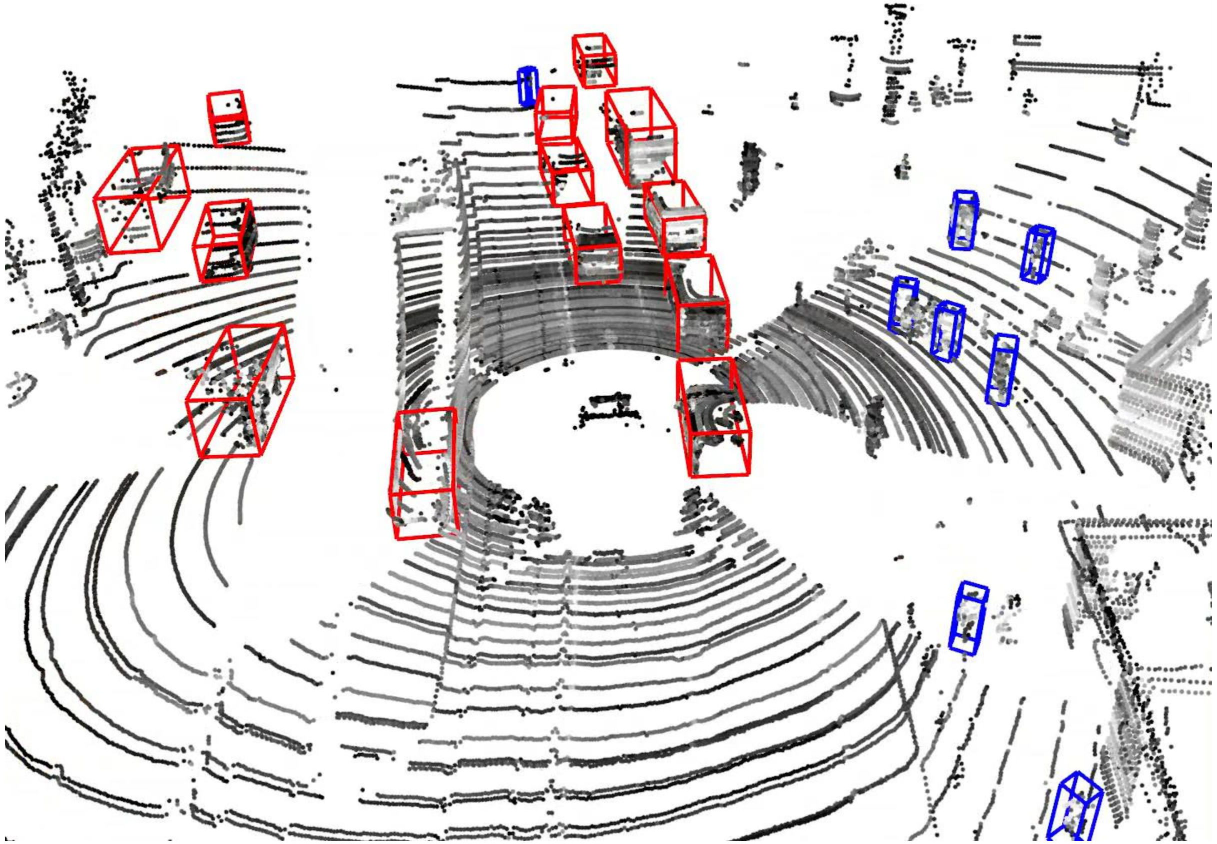
Low level change map for correcting object detection errors



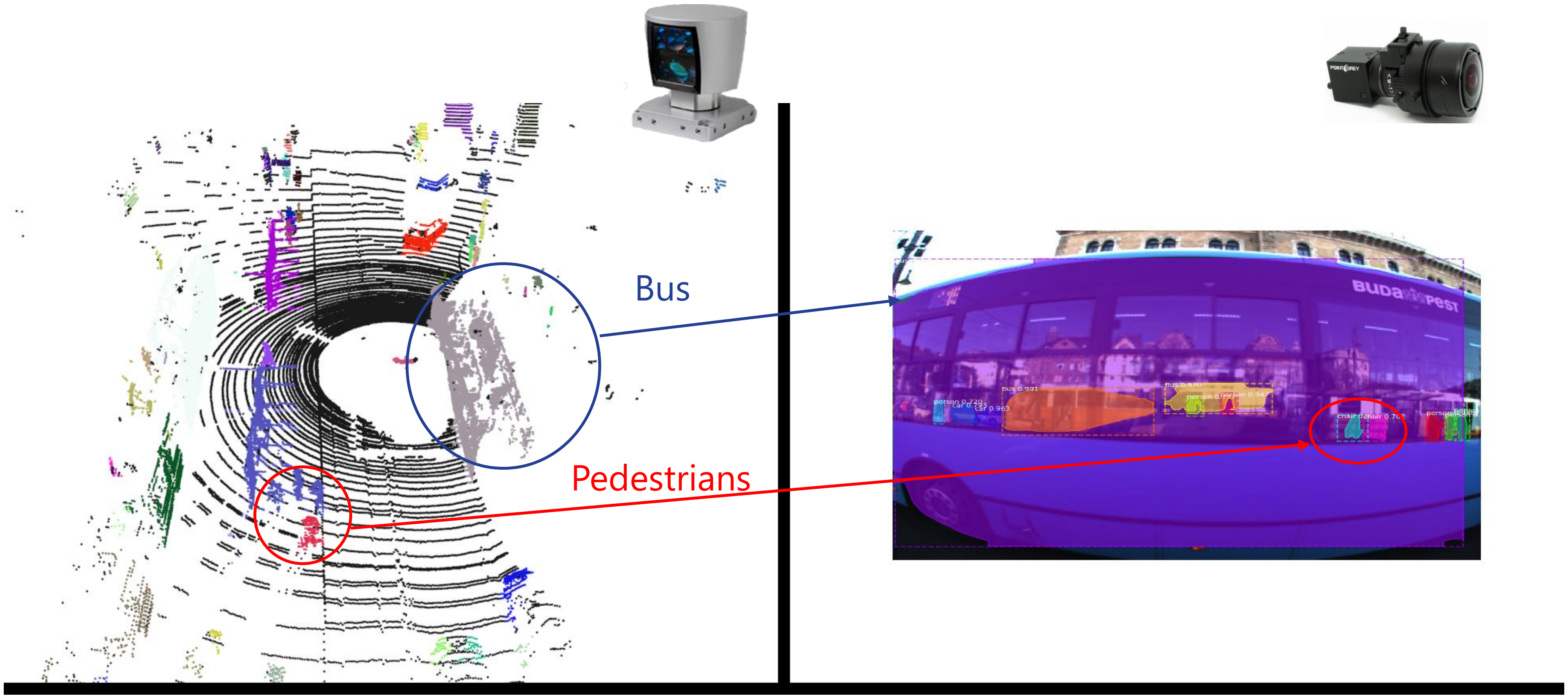
Ö. Zováthi, B. Nagy and *Cs. Benedek*: "Exploitation of Dense MLS City Maps for 3D Object Detection", ***International Conference on Image Analysis and Recognition (ICIAR)***, Póvoa de Varzim, Portugal, (virtual conference), June 24-26, 2020, vol- 12131 of ***Lecture Notes in Computer Science***, pp. 393-403, Springer, 2020

So far: object detection based on a <u>single onboard sensor</u>





Why camera-Lidar fusion is necessary?



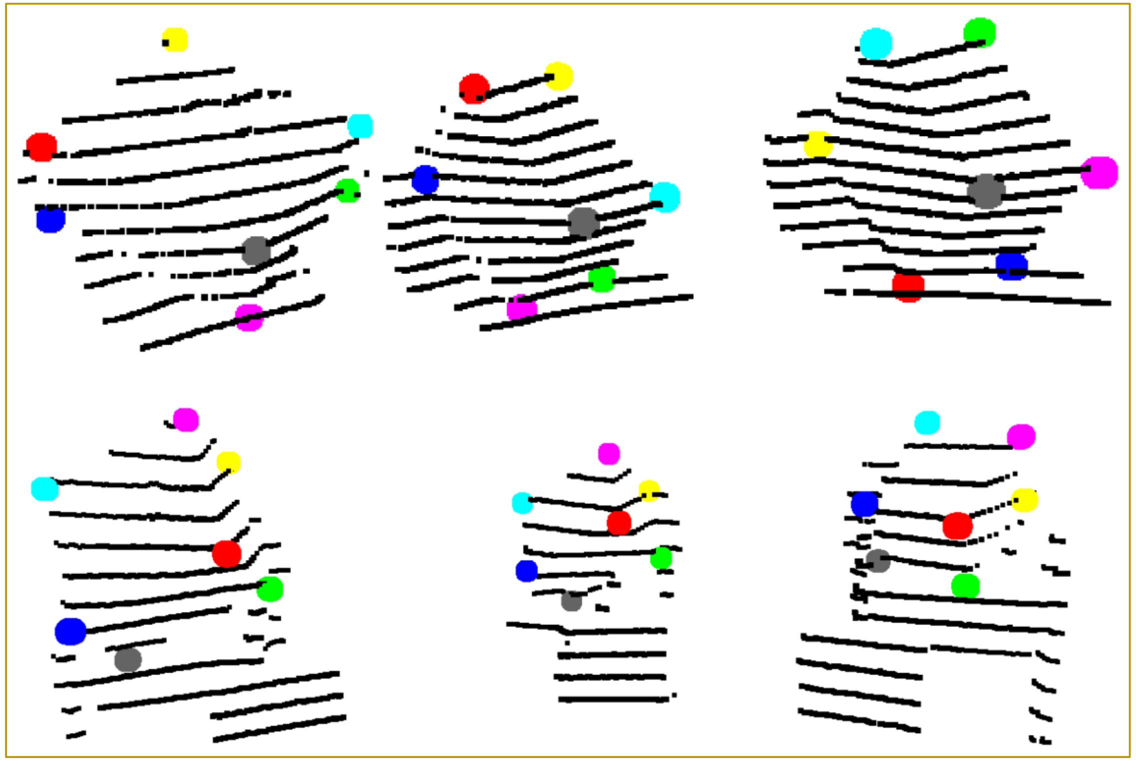
Lidar point cloud scene

– Low resolution: close objects (pedestrians) may be merged into the same blob

Camera image with detected objects

– *false pedestrian detections due to mirroring effects*

Target-based semi-manual reference method

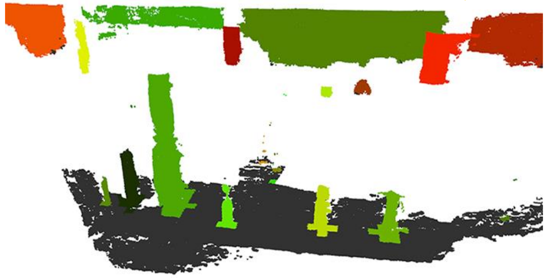


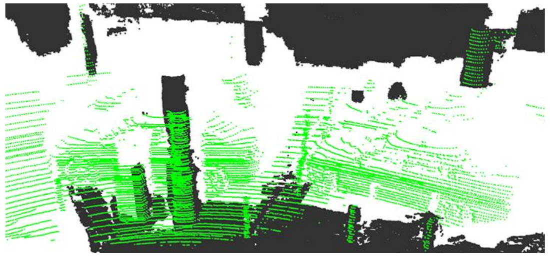
Pusztai, Z.; Eichhardt, I.; Hajder, L. Accurate Calibration of Multi-LiDAR-Multi-Camera Systems. Sensors ***2018****, 18, 2139.*

Automated Lidar-camera calibration
Working without calibration object

Structure from Motion (SfM)

SFM point cloud

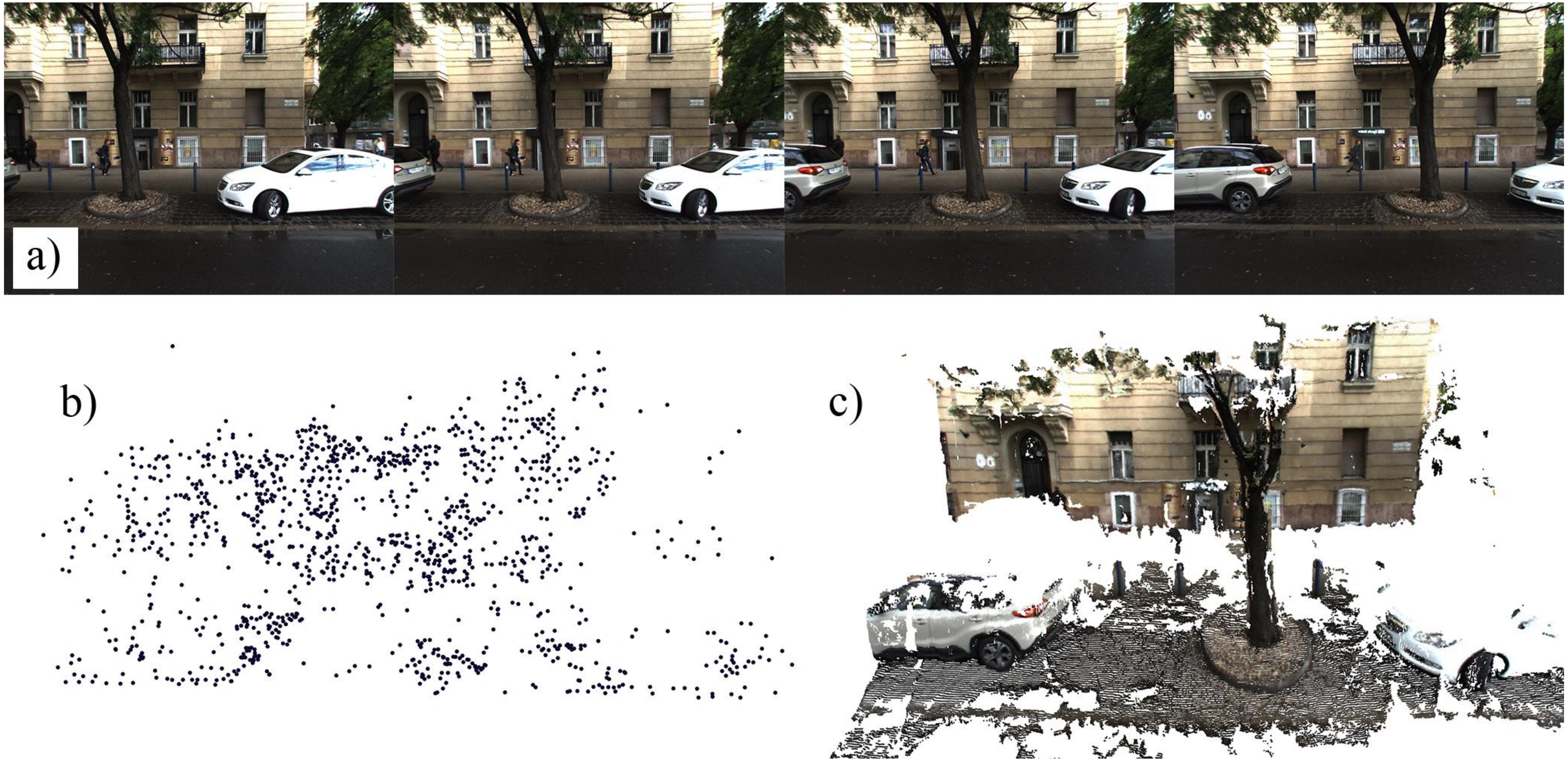




Camera images

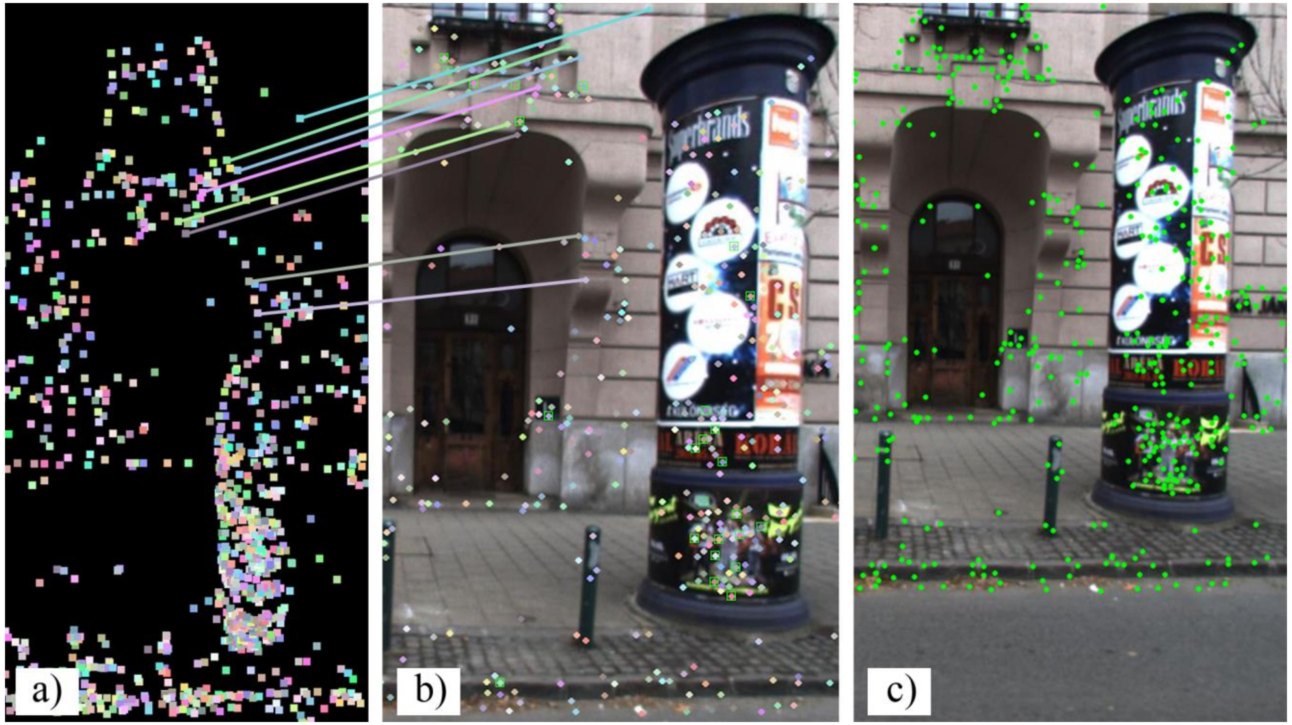
B. Nagy and Cs. Benedek: "On-the-Fly Camera and Lidar Calibration," Remote Sensing, vol 12, no. 7, article 1137, 2020, IF: 4.509

Structure from Motion (SfM)



Moulon, P.; Monasse, P.; Perrot, R.; Marlet, R. OpenMVG: Open Multiple View Geometry. Workshop on Reproducible Research in Pattern Recognition, 2016, pp. 60–74.

Structure from Motion (SfM)

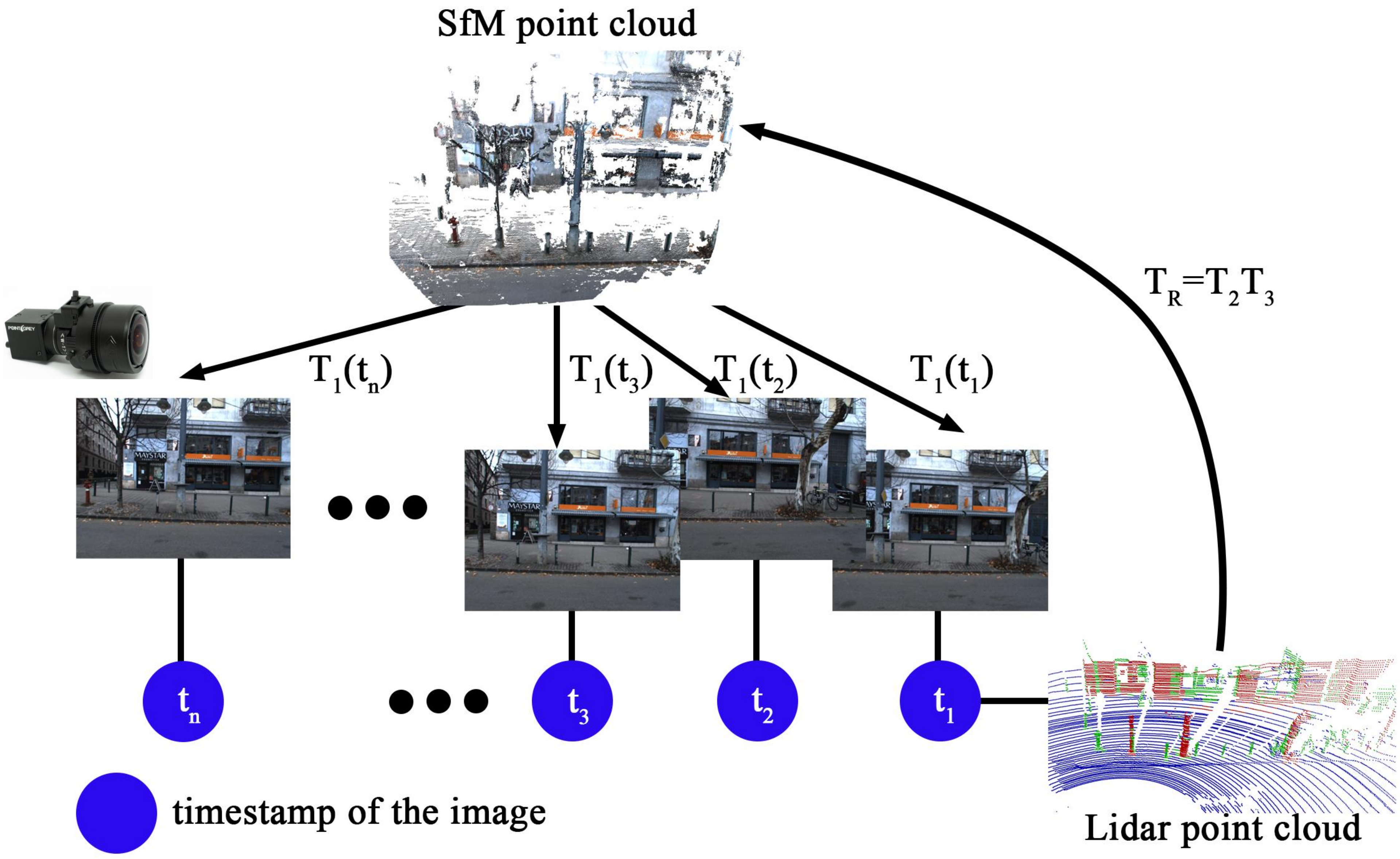


Reprojection calculation during the SfM method

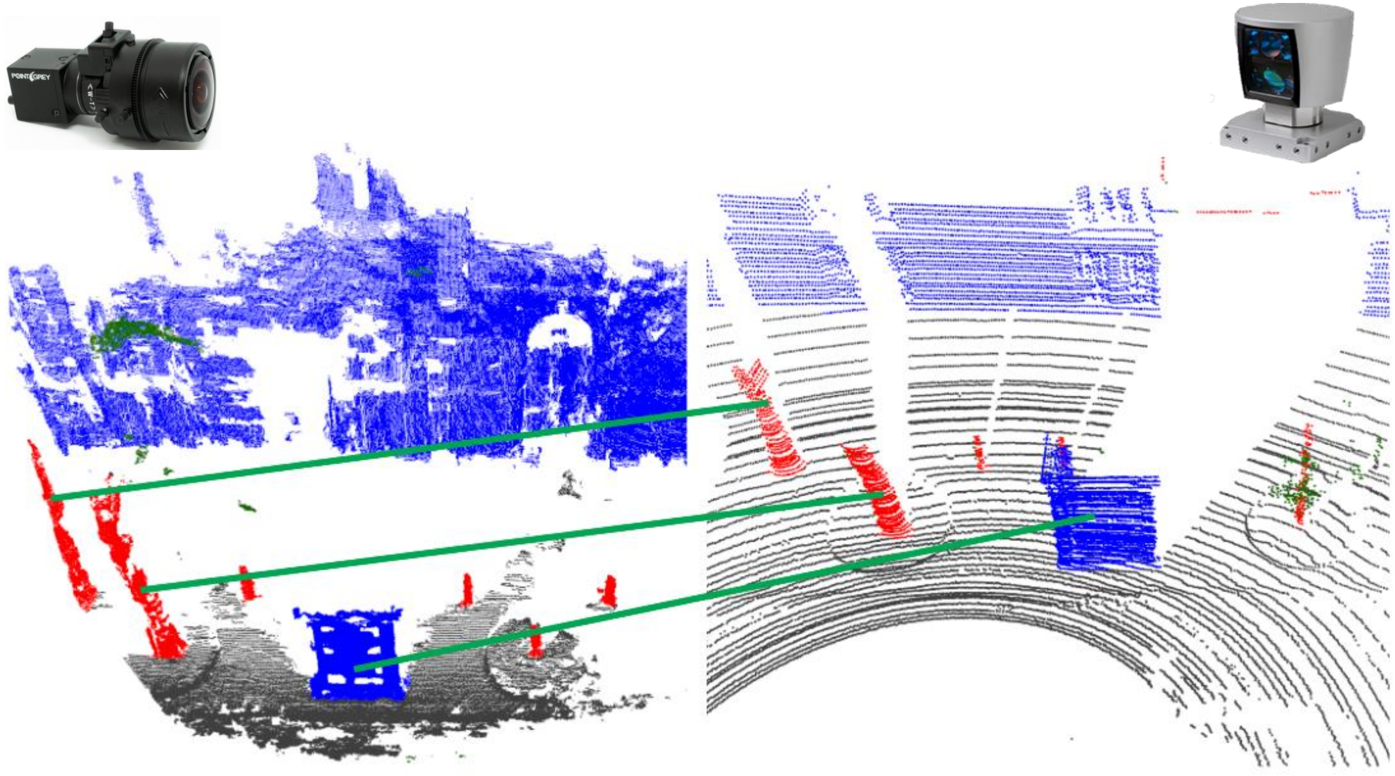
$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Perspective-*n*-Point (PnP) algorithm

Proposed method



Object based registration

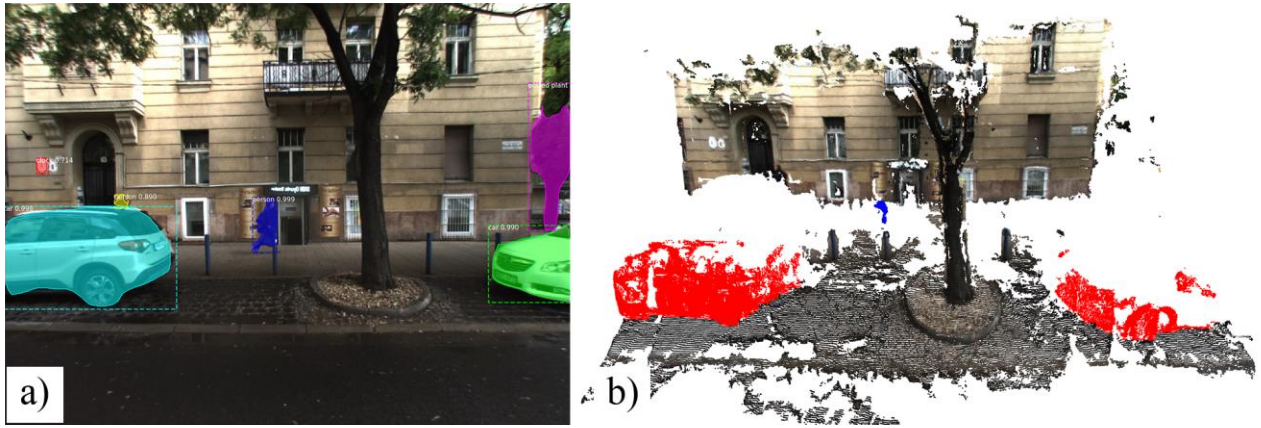


SFM point cloud   Lidar point cloud

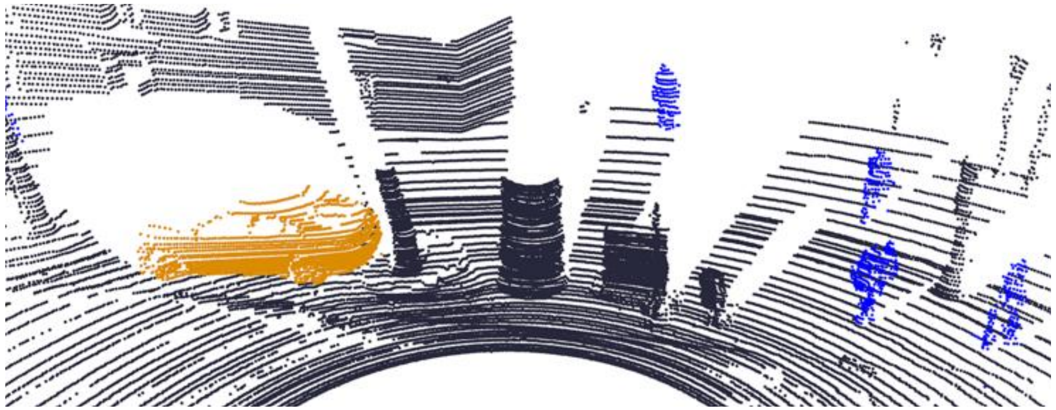
```
procedure Alignment(F1, F2, T)
 C1 <- ObjectDetect(F1)
 C2 <- ObjectDetect(F2)
 Initialize 4D accumulator array A
 for all o1 in C1 do
  for all o2 in C2 do
   for a in [0, 359] do
    o1’ <- Rot(a) * o1
    (dx, dy, dz) <- o2 – o1’
    A[dx,dy,dz,a] <- A[dx,dy,dz,a] + 1
   end for
  end for
 end for
 a, dx, dy, dz <- FindMax(A)
 F1, T <- TransformCloud(F1, a, dx, dy, dz)
end procedure
```

Using object detection for false alarm reduction

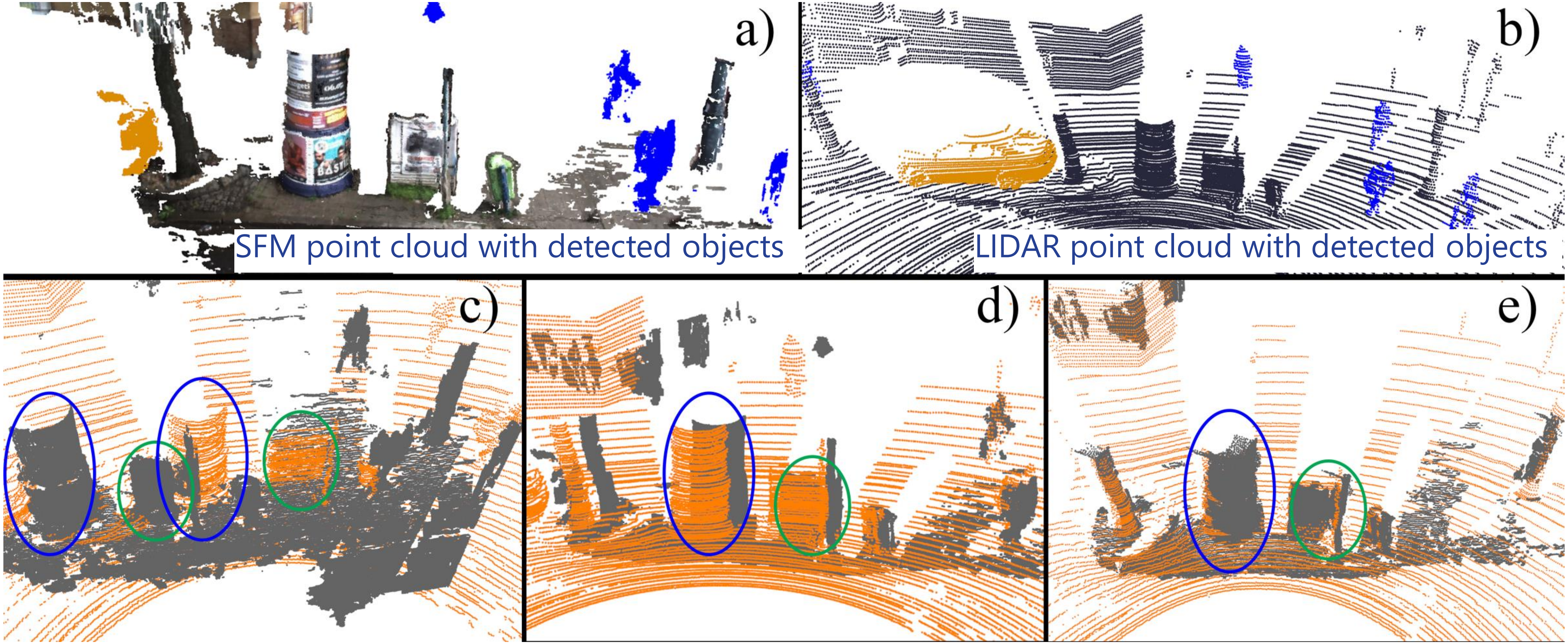
- Vehicle and pedestrian regions detected and excluded from the object based alignment step
- Camera data: 2D object detection by Mask R-CNN, then reprojection of results to SFM point cloud



- Lidar data: PointPillars based 3D object detection



Effect of dynamic object removal on object based registration



SFM point cloud with detected objects

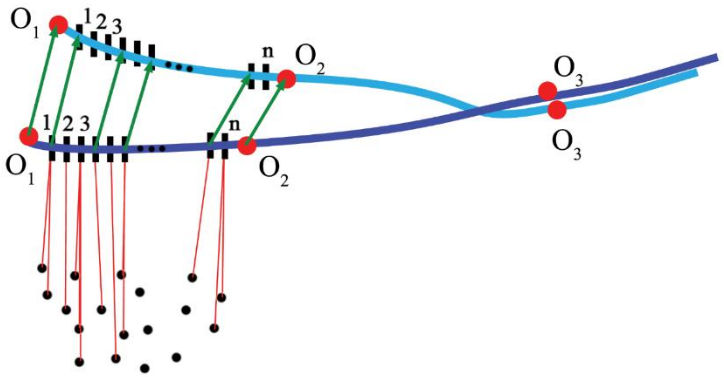
LIDAR point cloud with detected objects

Initial alignment

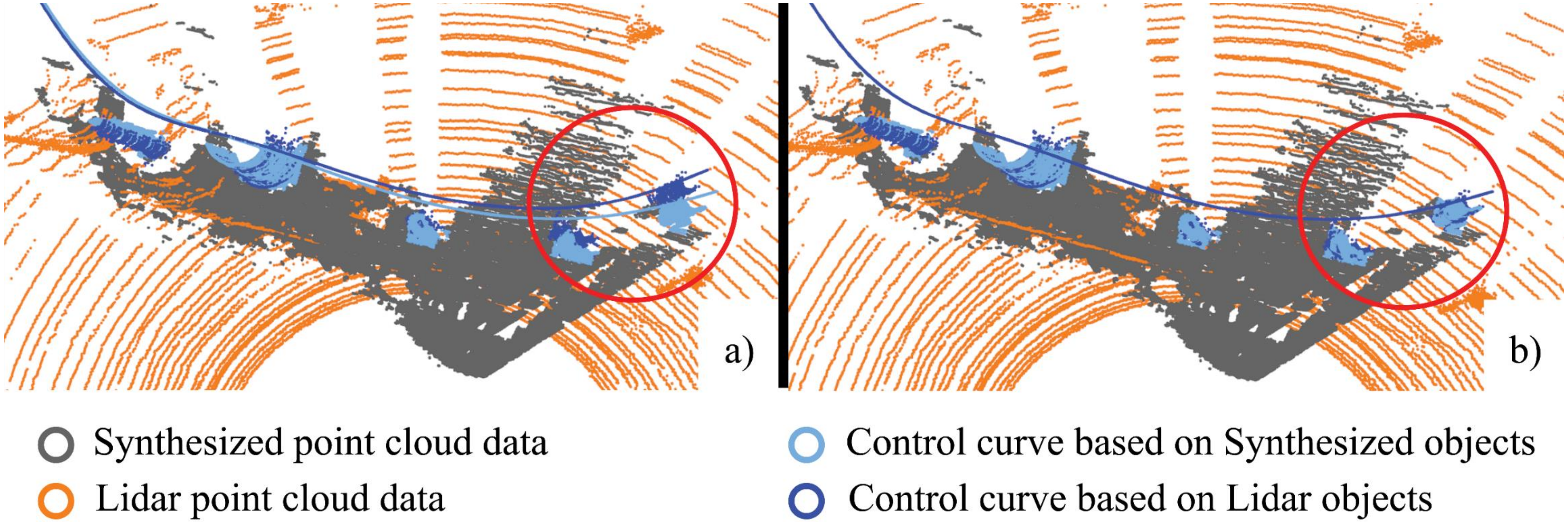
Matching without dynamic object removal

Matching with dynamic object removal

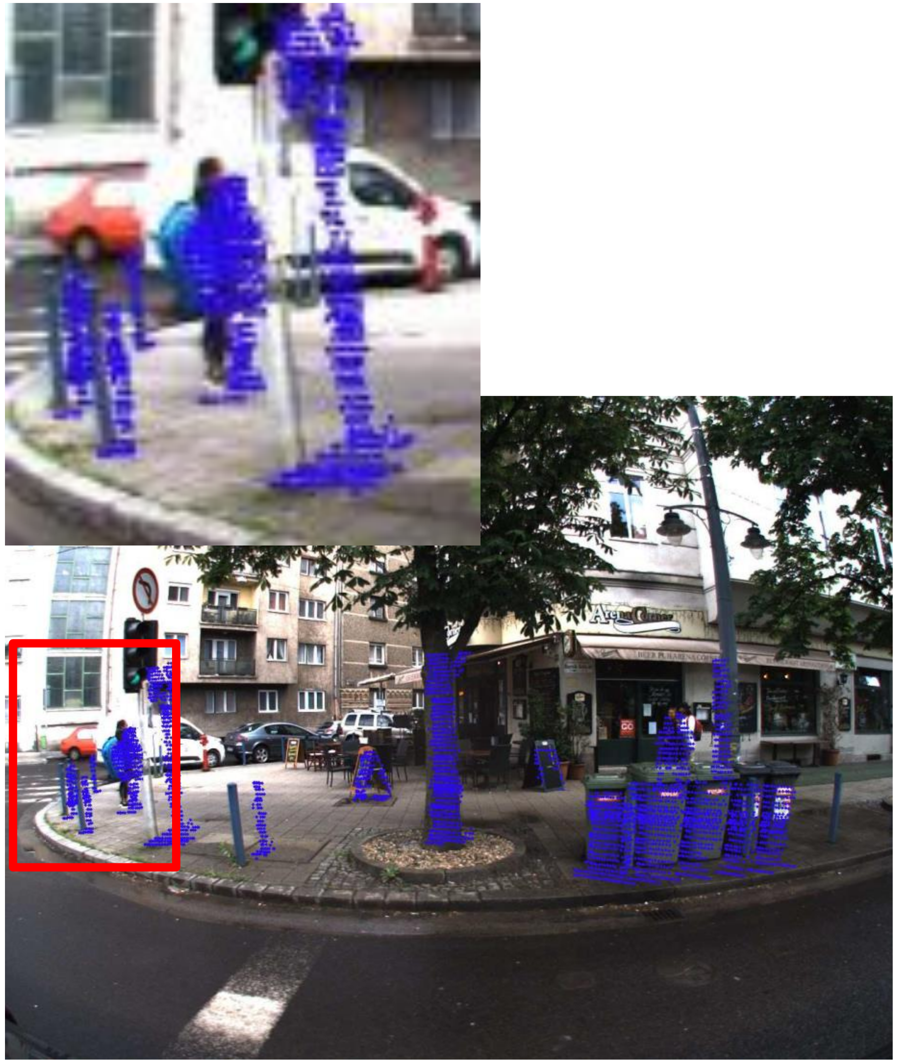
Curve based registration refinement



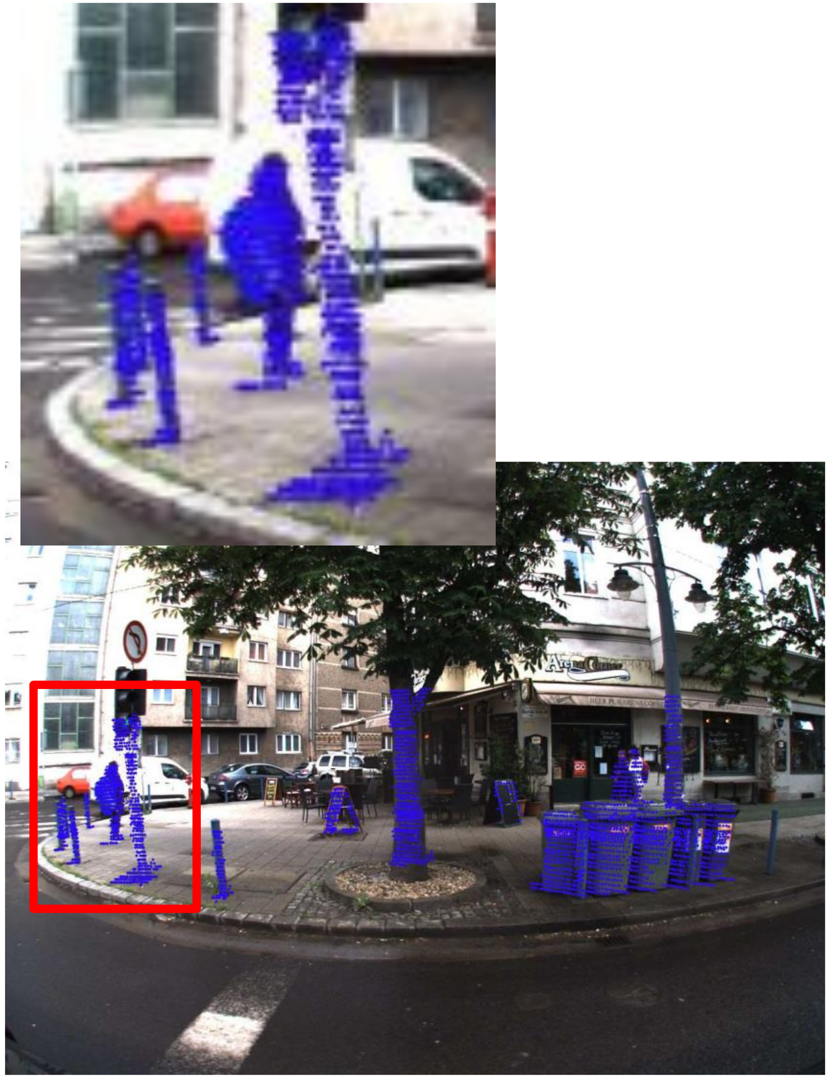
- Limitations of rigid body transform estimation
  - SFM pipeline artifacts - *deformations* due to *invalid depth calculation*
  - Issues with Lidar point clouds: *shape distortion* due to *vehicle motion*
- Control point based non-linear registration refinement
  - NURBS curves fit on the extracted object centers



Curve based registration refinement

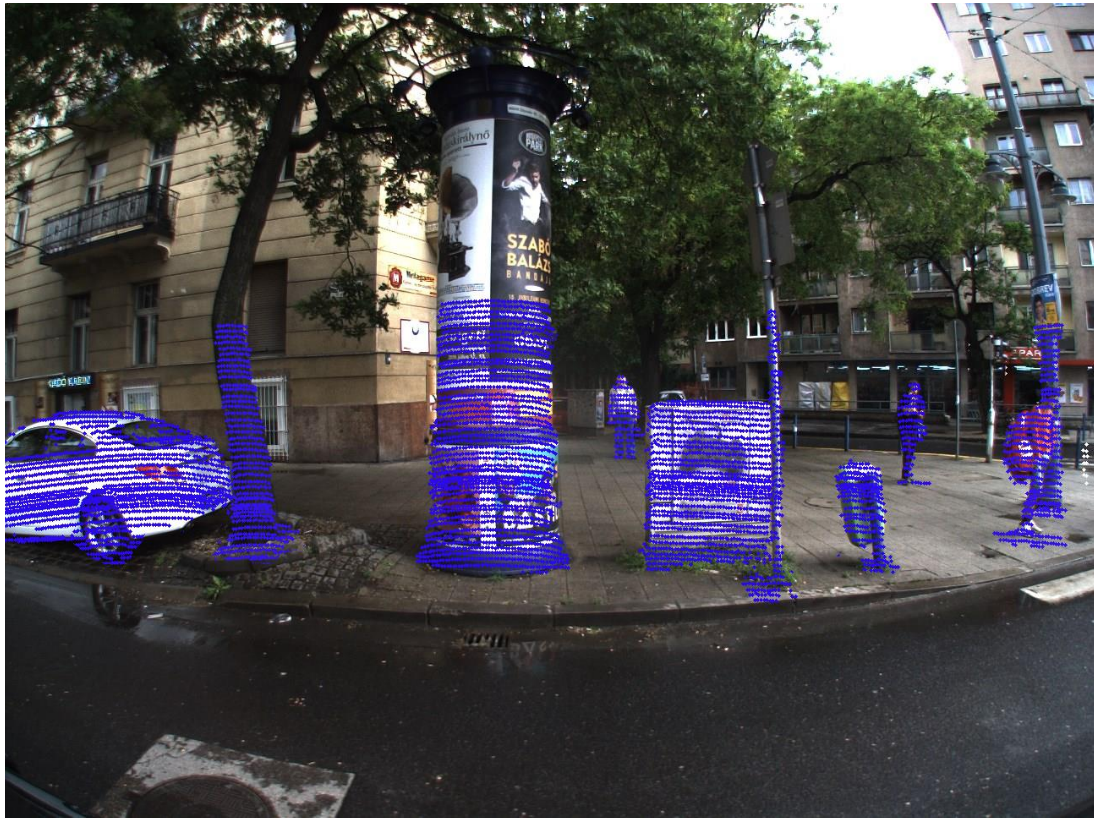


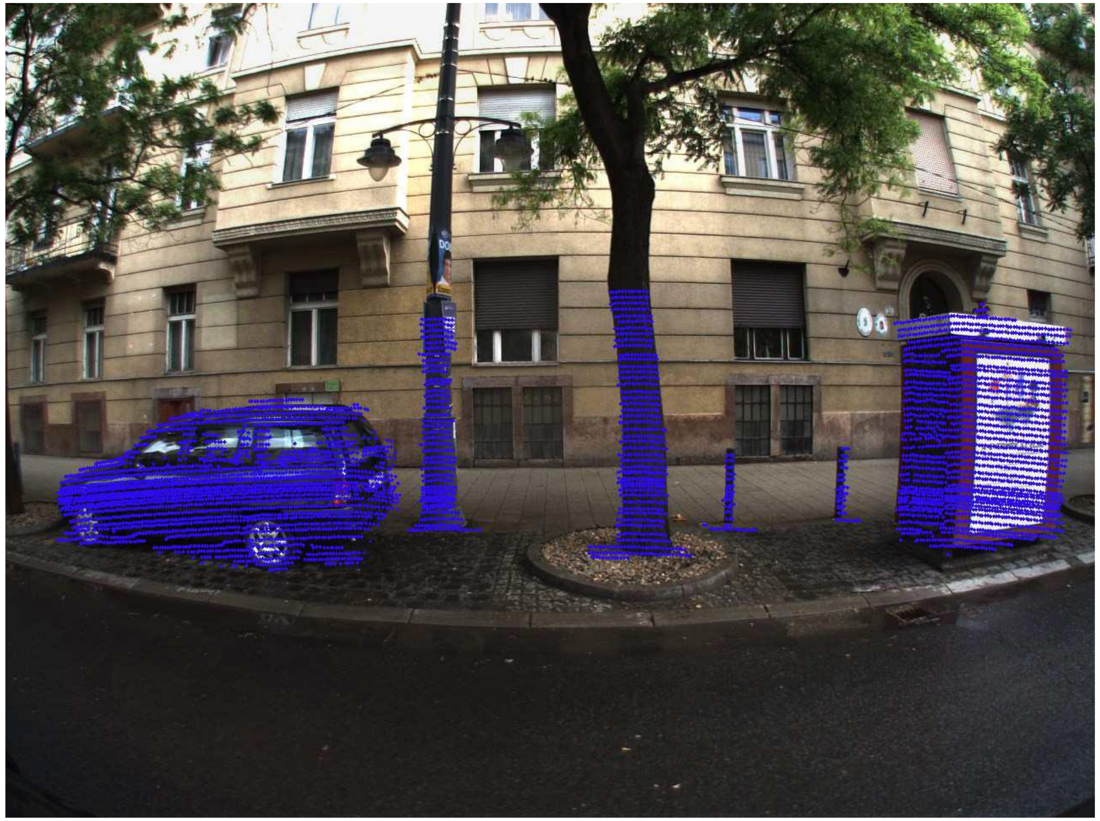
Rigid transfrom based registration



Result of non-linear refinement

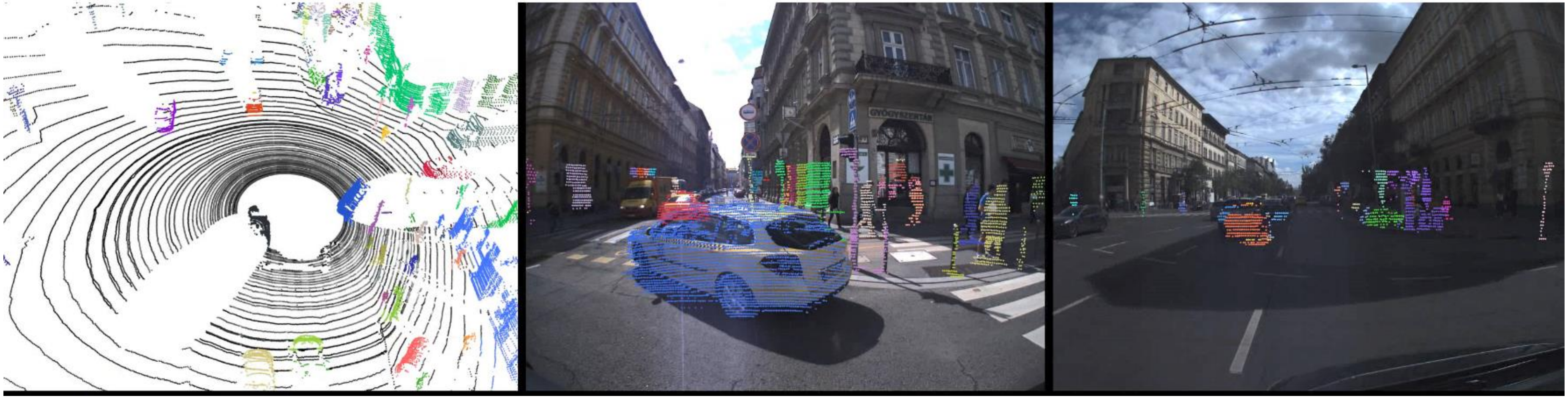
Some further results





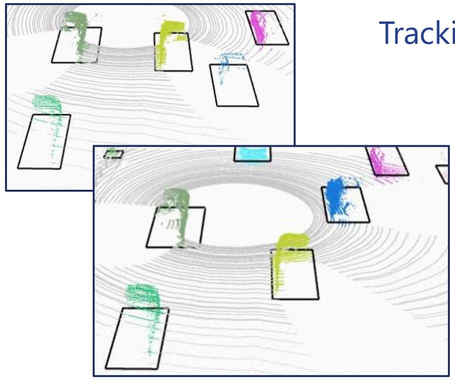
B. Nagy and *Cs. Benedek*: "On-the-Fly Camera and Lidar Calibration," ***Remote Sensing***, vol 12, no. 7, article 1137, 2020, IF: 4.509*

Camera-LiDAR fusion result 1

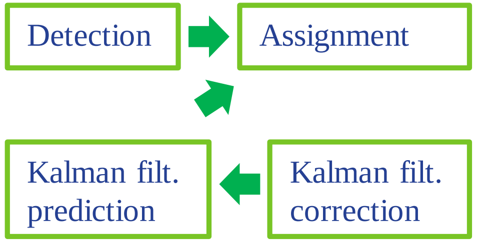


Steps:

1. Camera-LiDAR self calibration
2. Detection and tracking LiDAR point clouds
   - Fully geometry based object detection
   - Kalman-filter-based motion model for tracking
3. Projection of the object point clouds to the camera frames



Tracking: state machine



Thank you for your attention

- Contact: Csaba Benedek, benedek.csaba@sztaki.hu
- Contributors: Balázs Nagy, Örkény Zováthi

- Website:
  - Institute for Computer Science and Control, Machine Perception Research Laboratory https://www.sztaki.hu/en/science/departments/mplab
  - Research Group on Geo-Information Computing: http://mplab.sztaki.hu/geocomp/

- Funding:
  - Hungarian National Research, Development and Innovation Fund, under grants NKFIA K-120233 and 2018-2.1.3-EUREKA-2018-00032;
  - Michelberger Master Award of the Hungarian Academy of Engineering
  - European Union and the Hungarian Government from the projects EFOP-3.6.2-16-2017-00013 and EFOP-3.6.2-16-2017-00015