3-D Urban Objects Detection and Classification From Point Clouds

Yassin Alkhalili, Manisha Luthra, Amr Rizk, Boris Koldehofe
KOM - Multimedia Communications Lab
firstname.lastname@kom.tu-darmstadt.de

TECHNISCHE UNIVERSITÄT DARMSTADT

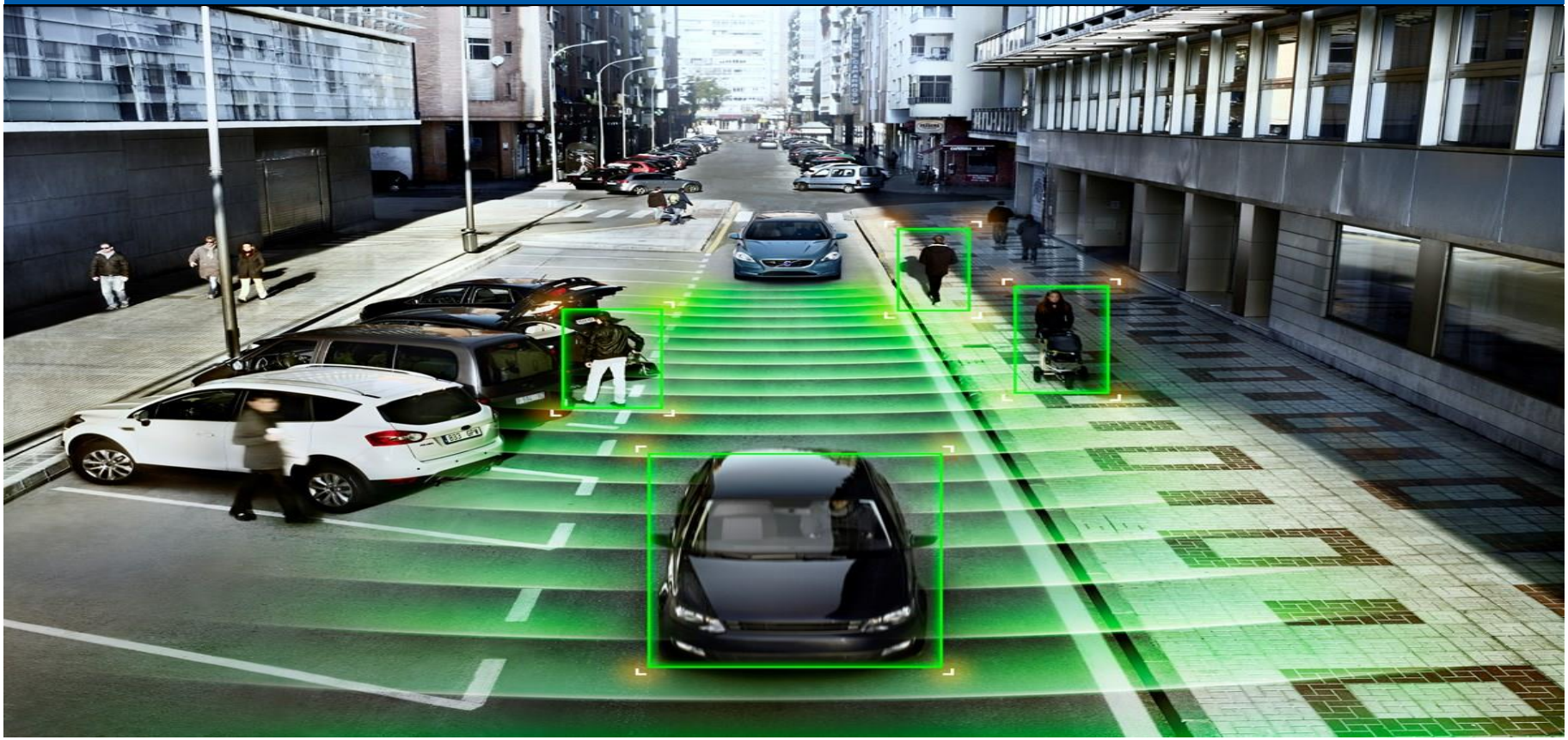


Image Source: https://www.ecofleetuk.com/advanced-road-safety-technology-to-be-standard-on-new-cars/

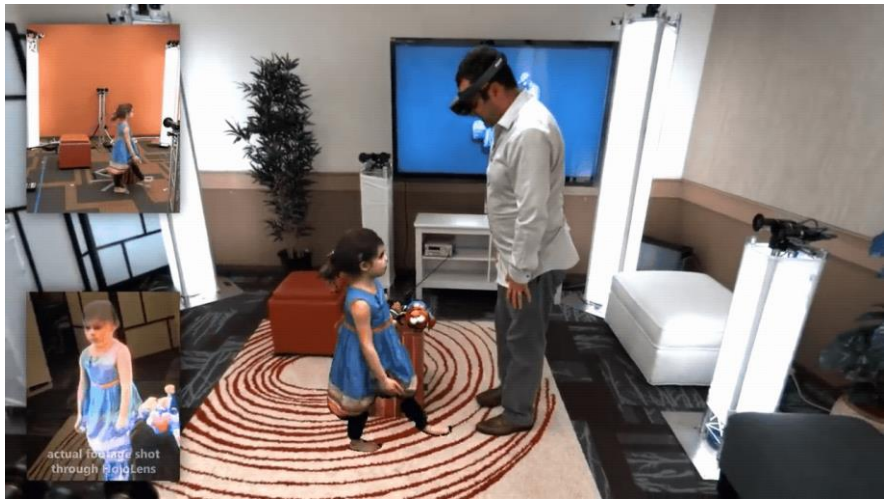
About me

TECHNISCHE UNIVERSITÄT DARMSTADT

- Phd student at **KOM** - **Multimedia Communications Lab** – Tu-Darmstadt
- I am part of "**MAKI– Multi-Mechanisms Adaptation for the Future Internet**"
- Research Interest: **Point cloud Processing and Streaming**

SFB 1053

Immersive Communications



Department of Electrical Engineering
and Information Technology
Multimedia Communications Lab - KOM

TECHNISCHE UNIVERSITÄT DARMSTADT

Yassin Alkhalili, M.Sc.

Yassin.Alkhalili@KOM.tu-darmstadt.de
Rundeturmstr. 10
64283 Darmstadt/Germany
www.kom.tu-darmstadt.de

Phone +49 6151 16-20491
Fax +49 6151 16-29109

Image source: https://youtu.be/7d59O6cfaM0

How to define a Point Cloud?

- **A collection of points in 3D space**

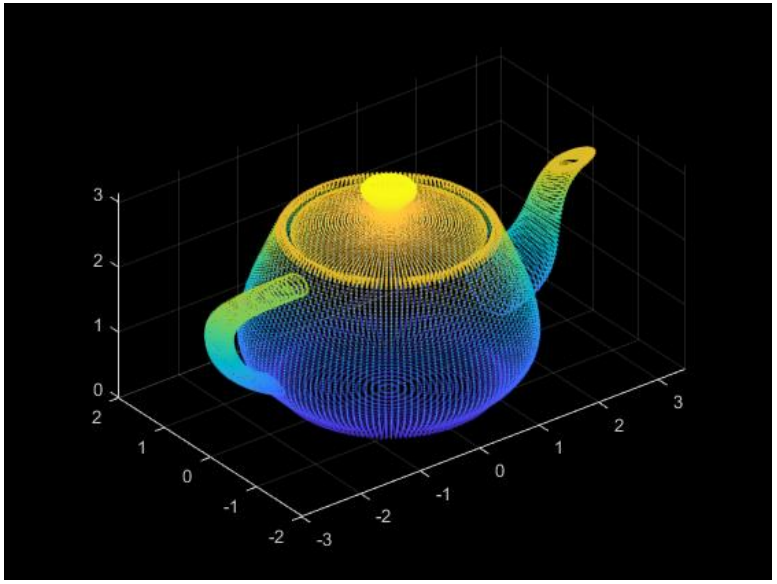
Unbenannt - Editor

Datei Bearbeiten Format Ansicht ?

```
ply
format ascii 1.0
comment Version 2, Copyright 2017, 8i Labs, Inc.
comment frame_to_world_scale 0.179523
comment frame_to_world_translation -45.2095 7.18301 -54.3561
comment width 1023
element vertex 879340
property float x
property float y
property float z
property uchar red
property uchar green
property uchar blue
end_header
171 62 251 175 154 140
171 63 251 172 150 135
171 61 252 174 153 139
171 61 253 172 151 138
170 61 255 152 136 124
171 61 254 166 147 134
171 61 255 160 141 129
170 62 253 171 151 139
170 63 252 173 153 140
170 63 253 175 155 142
171 62 252 175 153 140
171 62 253 172 152 139
171 63 252 174 153 139
171 63 253 176 156 144
170 62 254 165 146 134
170 62 255 159 142 130
170 63 254 171 152 140
```

← Number of points

One point: X, Y, Z, R,G,B



Point cloud format (Example)

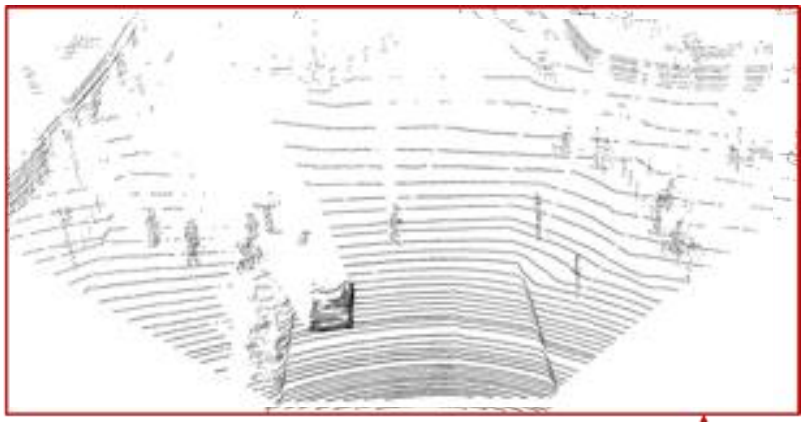
Image source: https://es.mathworks.com/help/vision/ref/pcwrite.html

Motivation

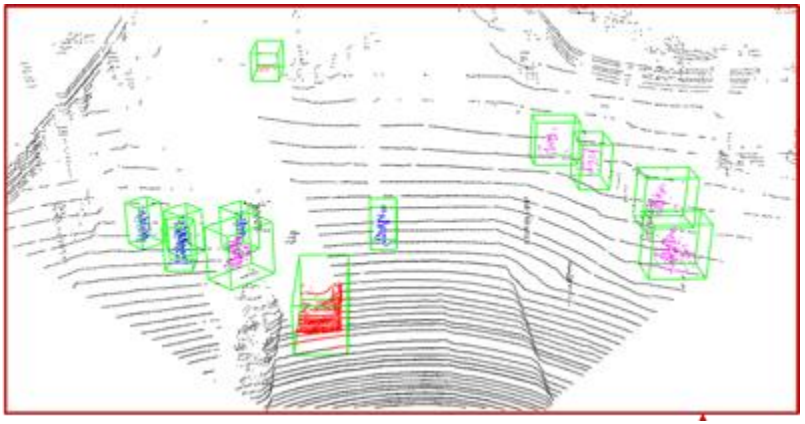
TECHNISCHE UNIVERSITÄT DARMSTADT

- **Point Cloud can be created by LiDAR**
  - Easy to produce

- **LiDAR VS RADAR and IR**
  - insensitivity to colors, lighting conditions, and radial distortion



Row Data

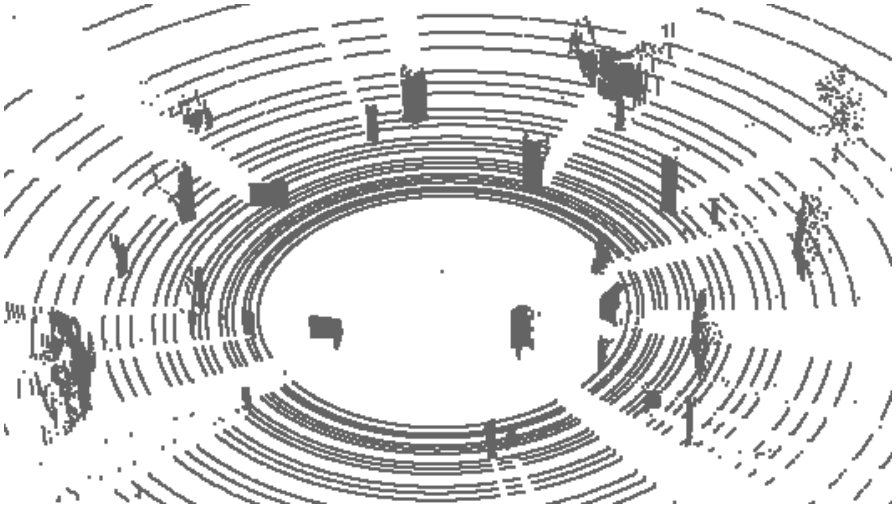


Knowledge about the environment

Image source: https://velodynelidar.com/hdl-32e.html
https://ori.ox.ac.uk/efficient-object-detection-from-3d-point-clouds/

DEBS19 – Grand Challenge

TECHNISCHE UNIVERSITÄT DARMSTADT



Object Type	Count
Pedestrian	3
ATM	1
Phone Booth	2

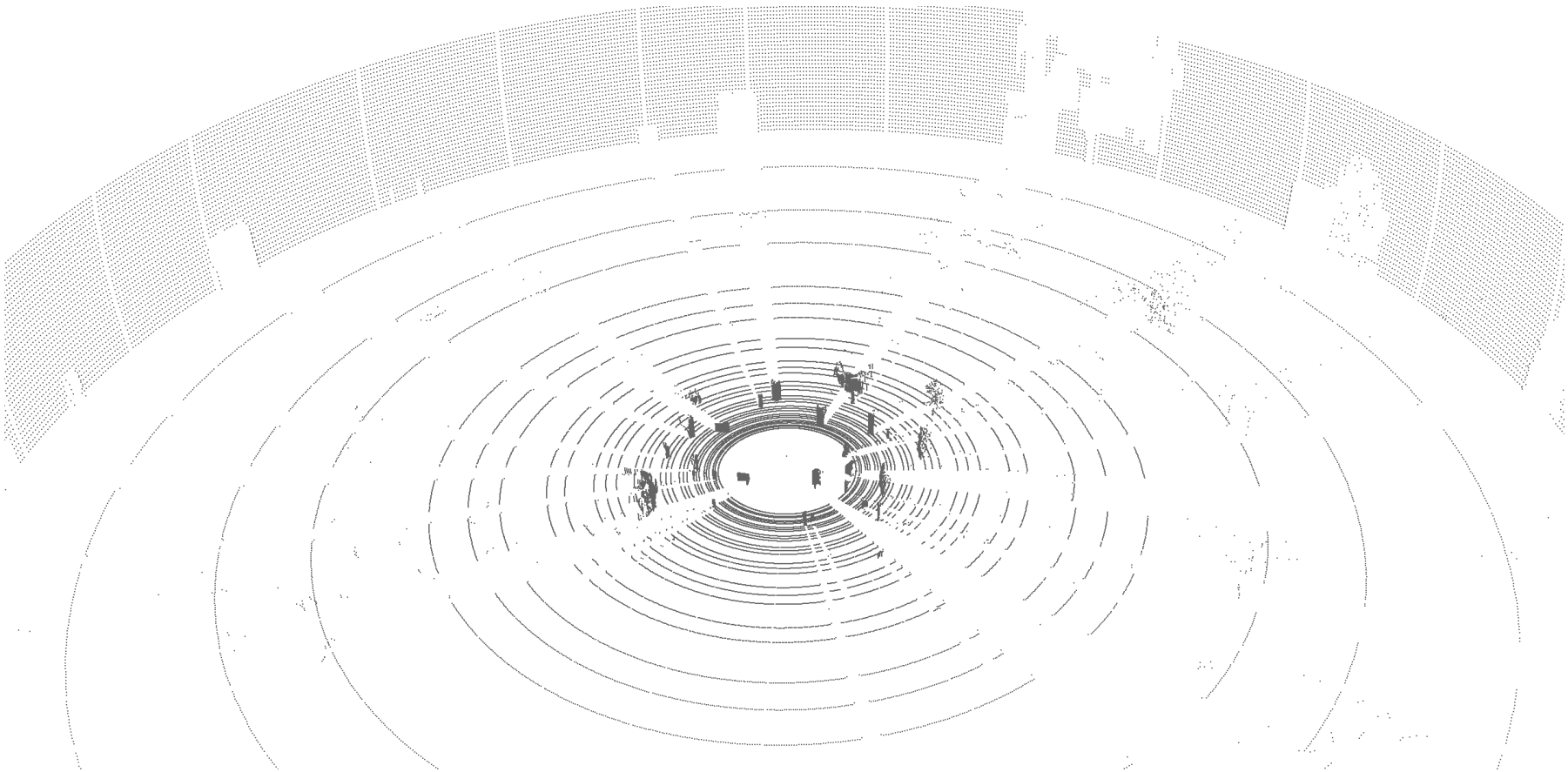
Input

output

This work addresses the problem of 3D Urban object class recognition in different LiDAR scenes

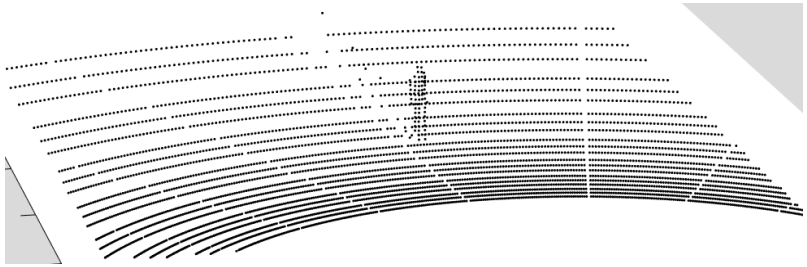
Challenge Dataset

TECHNISCHE UNIVERSITÄT DARMSTADT

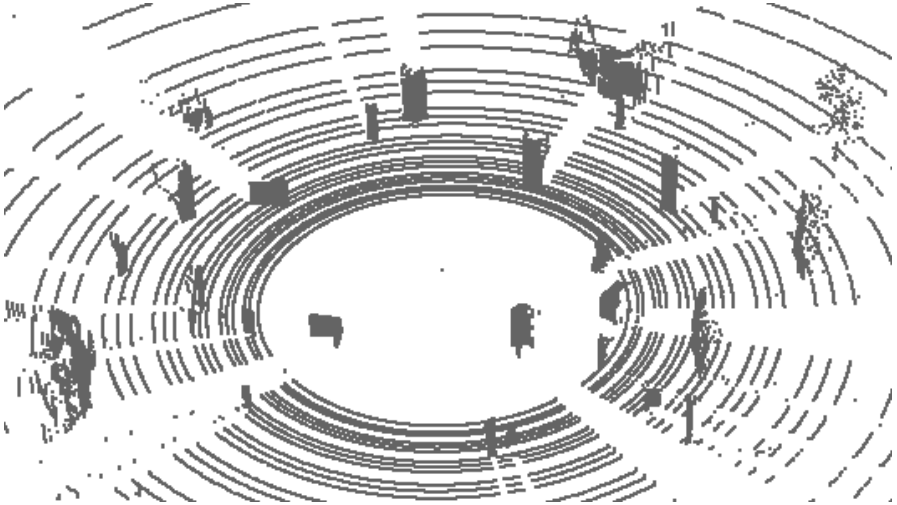


Bird’s eye view of a LiDAR scene

Challenge Dataset



Single Object Scene



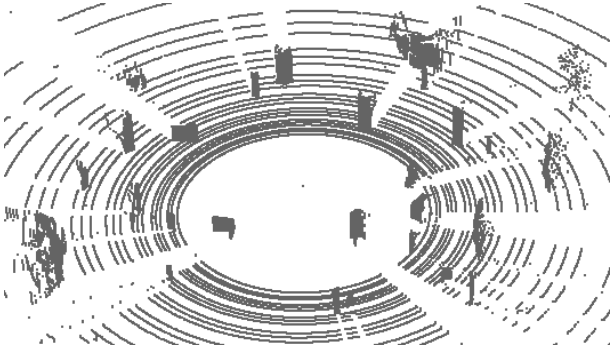
Multiple Objects Scene

Challenge Dataset

TECHNISCHE UNIVERSITÄT DARMSTADT

- **The dataset includes 28 different object classes**
- **50 single object scene per object class → 1400 single object scenes**
- **600 multiple object scenes**
- **Each multiple object scenes contains 10 to 50 objects**

- **Labels are assigned to the entire scene and not the individual objects.**



Scene

[ATM, Car, Pedestrian]

Labels

Challenges in the Dataset

TECHNISCHE UNIVERSITÄT DARMSTADT

- **High number of objects per scene (10-50)**
- **Objects are in different distance to the LiDAR → different object scales and different point densities**
- **Possibly interwoven objects**
- **Object classes are very similar to each other**

BMW

Toyota

- **The amount of training data 600 multiple objects scenes are limited**

Method

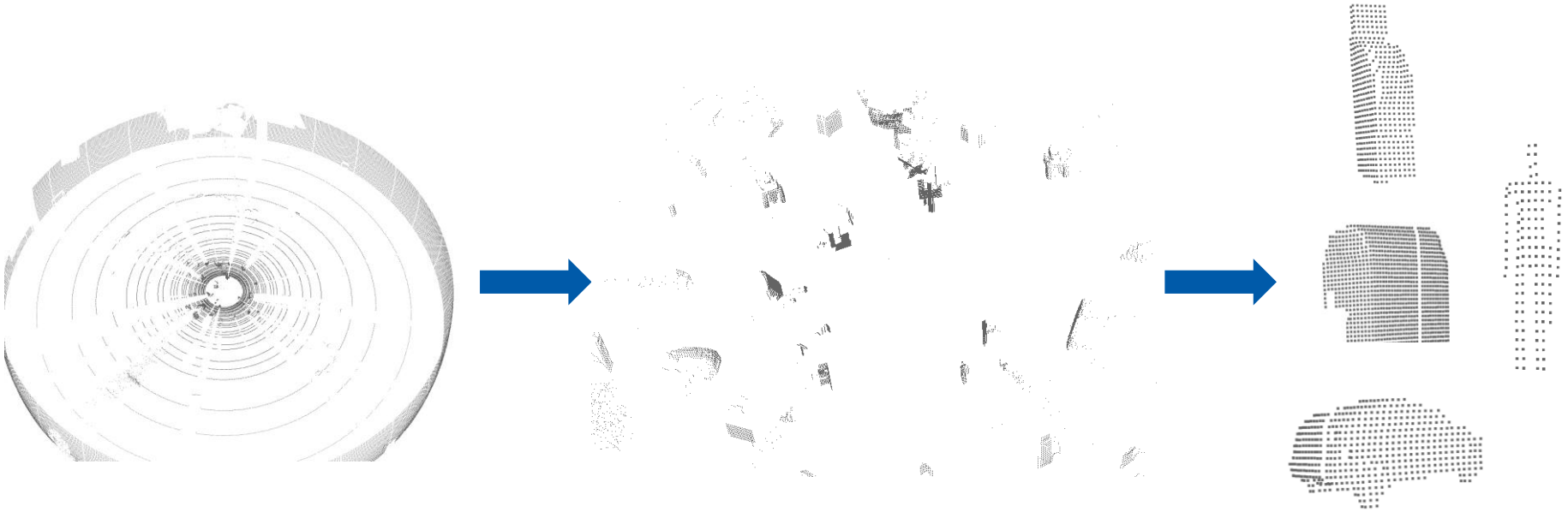
TECHNISCHE UNIVERSITÄT DARMSTADT

Goal: learn object types and recognize them in low latency

Proposed method:

1. Detecting candidate objects in a given scene
2. Generating features for each object
3. Use classification model for training and predicting

1- Detecting Candidate Objects in a Given Scene



Given Scene

Removing walls and floors

X and Z ∈ [-25,25]
Y > -2.5

Clean Scene – 60% data less

Euclidean Segmentation

```
clustering.setClusterTolerance(0.3); // 30 cm
clustering.setMinClusterSize(50);
clustering.setMaxClusterSize(5000);
```

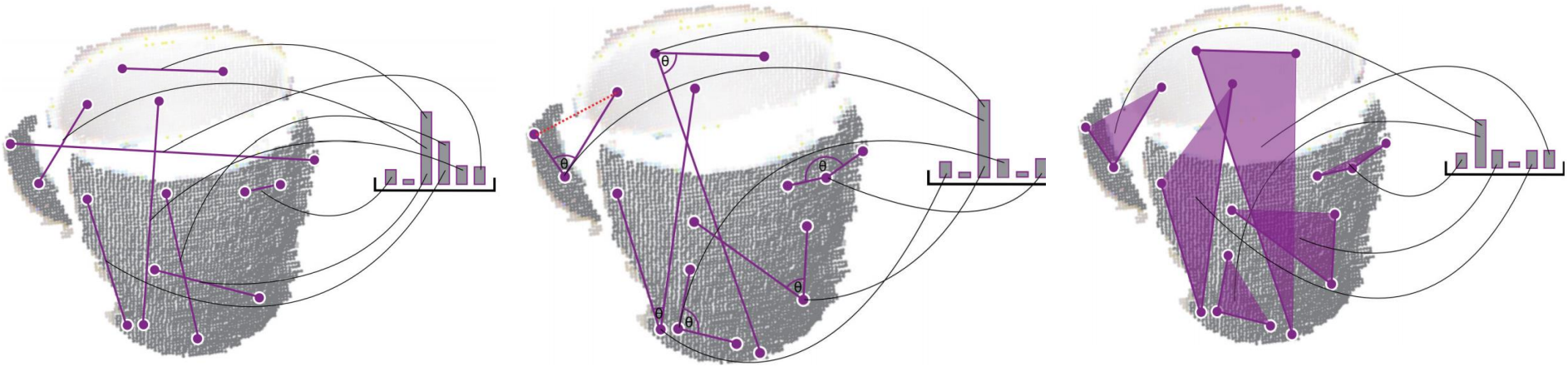
Objects

2- Object Features Generation
Ensemble of Shape Functions (ESF) [1]

TECHNISCHE UNIVERSITÄT DARMSTADT

ESF shape descriptor based on three distinct shape functions describing:

- Point Distance distribution
- Angle distributions
- Area distribution



1 - Wohlkinger and M. Vincze, "Ensemble of shape functions for 3D object classification," 2011 IEEE International Conference on Robotics and Biomimetics, Karon Beach, Phuket, 2011, pp. 2987-2992.

ESF Descriptor

TECHNISCHE UNIVERSITÄT DARMSTADT

Why ESF [1]:

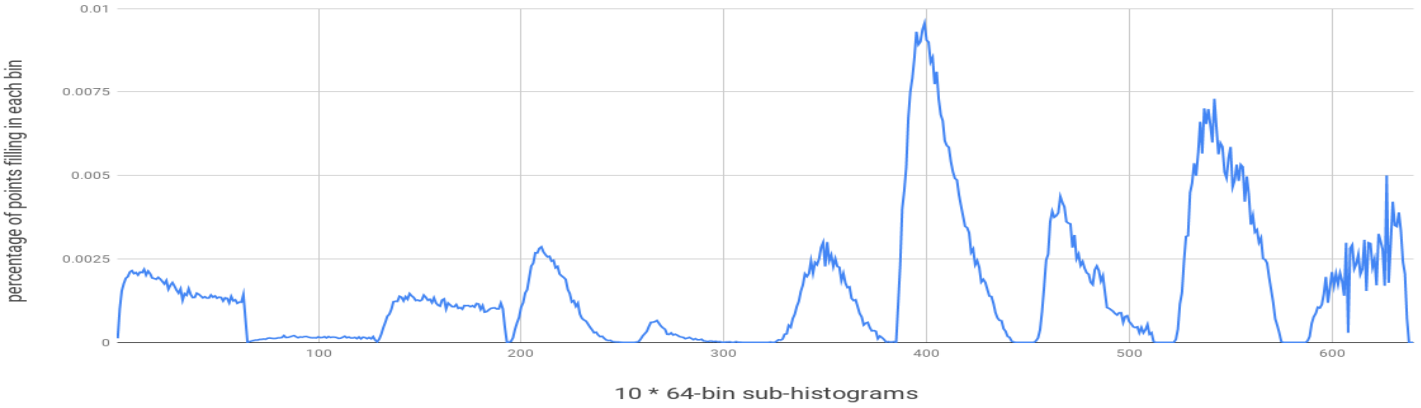
- Scale to hundreds of object types
- No preprocessing necessary e.g. surface normal calculation
- Handles data error such as outliers holes noise gracefully

1 - Wohlkinger and M. Vincze, "Ensemble of shape functions for 3D object classification," 2011 IEEE International Conference on Robotics and Biomimetics, Karon Beach, Phuket, 2011, pp. 2987-2992.

2- Object Features Generation

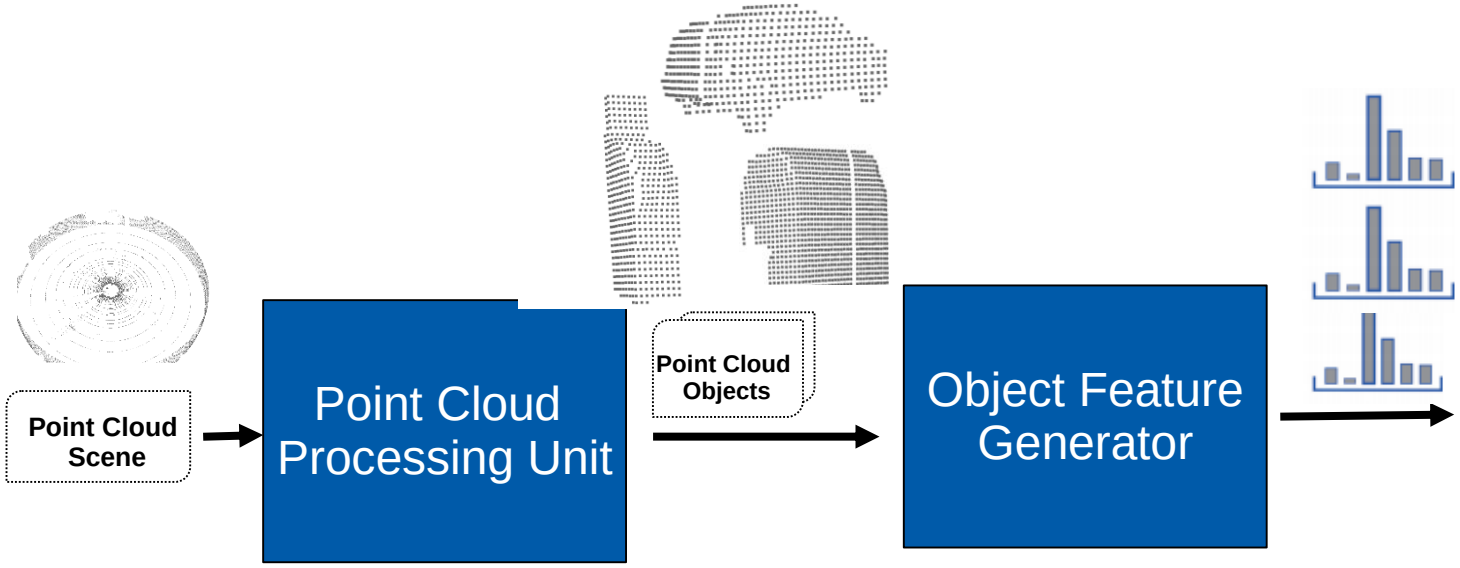
ESF

BmwX5Simple



0.662	1.987	1.987	1.325	1.325	1.987	1.325	1.325	1.987	3.974	1.325

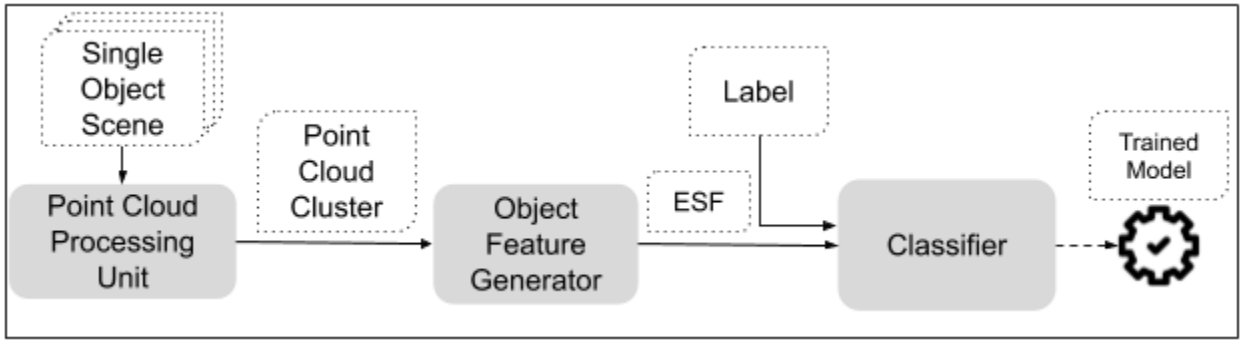
System Model



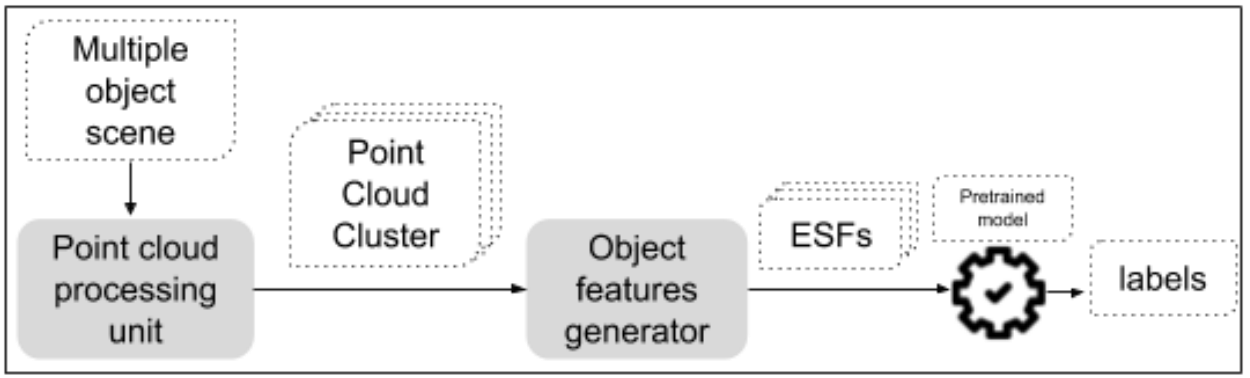
Training vs Predicting

TECHNISCHE UNIVERSITÄT DARMSTADT

Training



Predicting



Evaluation

TECHNISCHE UNIVERSITÄT DARMSTADT

➢ **We used single object scenes (labelled data) to train the model**

- Training Data:
  - 767 objects
  - 28 object types
- Testing Data:
  - 330 objects

Classifier	Accuracy	Precision	Recall
K Neighbors	0.48	0.50	0.49
Random Forest	**0.59**	**0.61**	**0.60**
Gradient Boosting	0.49	0.52	0.50

Evaluation

Class	Precision	Recall	Avg. points
Atm	0.67	0.60	245
Bench	0.47	0.69	172
BigSassafras	0.85	0.79	970
BmwX5Simple	0.42	0.62	874
ClothRecyclingContainer	0.38	0.40	437
Cypress	0.90	1.00	598
DrinkingFountain	**0.00**	**0.00**	**53**
ElectricalCabinet	0.45	0.50	101
EmergencyPhone	0.88	0.94	131
FireHydrant	**0.00**	**0.00**	**48**
GlassRecyclingContainer	0.88	0.88	417
IceFreezerContainer	0.36	0.45	270
Mailbox	0.22	0.29	90
MetallicTrash	1.00	0.25	69
MotorbikeSimple	0.54	0.64	136
Oak	0.78	0.78	410
OldBench	0.50	0.11	108
Pedestrian	0.50	0.43	112
PhoneBooth	0.79	0.69	467
PublicBin	0.47	0.80	81
Sassafras	0.50	0.56	640
ScooterSimple	0.67	0.91	150
ToyotaPriusSimple	0.73	0.69	728
Tractor	0.60	0.40	456
TrashBin	0.43	0.21	114
TrashContainer	0.56	0.33	338
UndergroundContainer	0.25	0.44	97
WorkTrashContainer	0.69	0.50	306

Evaluation

TECHNISCHE UNIVERSITÄT DARMSTADT

- **Performance of the our system as reported by DEBS'19 online evaluation system**

System	Accuracy	Precision	Recall	Runtime(sec)	#Scenes
AOC_KOM	0.479	0.56	**0.67**	557	442

System Limitations

TECHNISCHE UNIVERSITÄT DARMSTADT

- **Segmentation**
  - might divide the same objects into many
  - Very closed objects to each other might be considered one object

- **Execution time**
  - We develop our solution in Python, but ESF is available only in C++, so we had to call C++ from python which impose a delay which can be easily avoided if we implement the whole solution in C++

Questions

TECHNISCHE UNIVERSITÄT DARMSTADT

Department of Electrical Engineering
and Information Technology
Multimedia Communications Lab - KOM

TECHNISCHE UNIVERSITÄT DARMSTADT

Yassin Alkhalili, M.Sc.

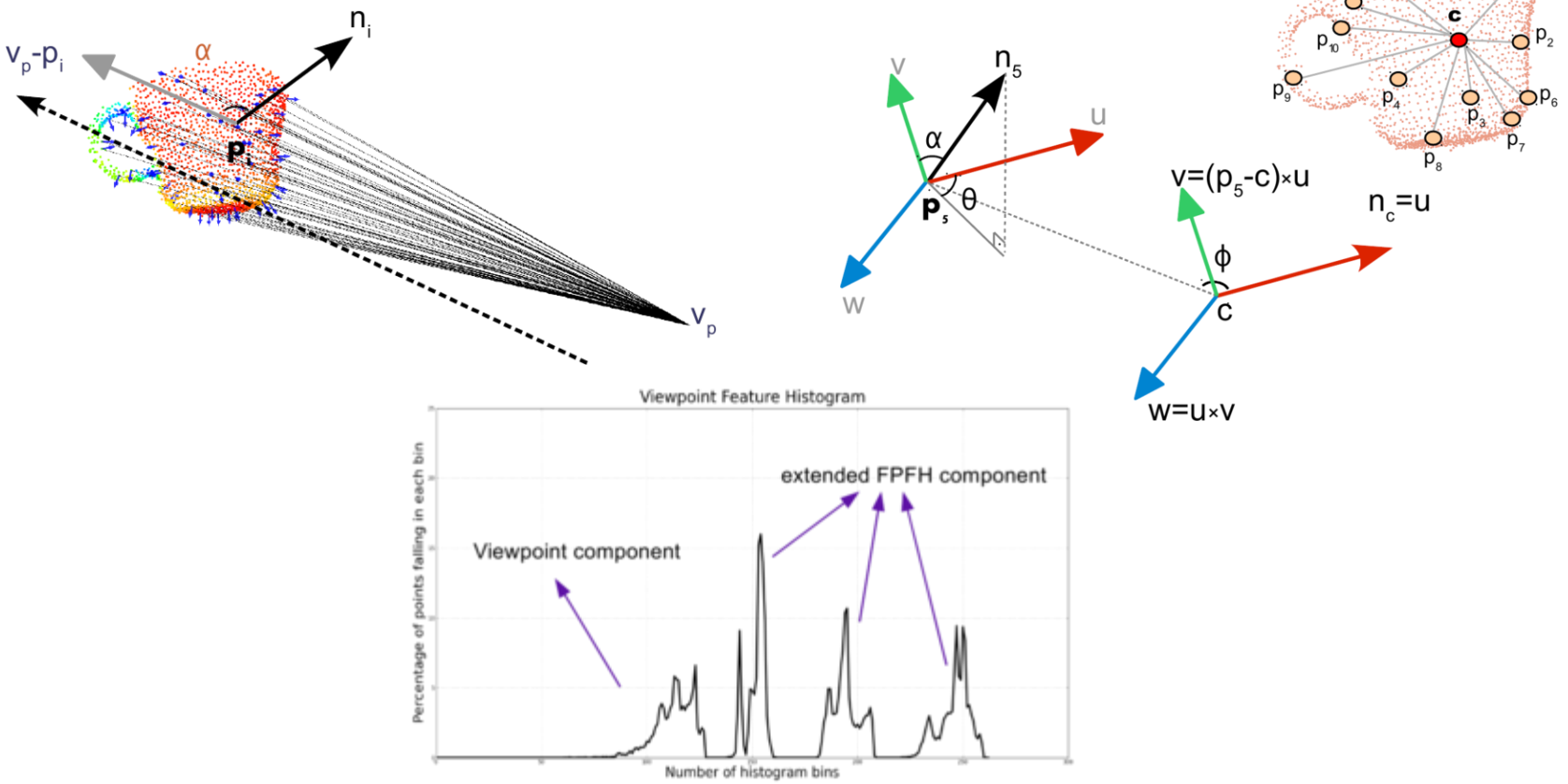
Yassin.Alkhalili@KOM.tu-darmstadt.de
Rundeturmstr. 10
64283 Darmstadt/Germany
www.kom.tu-darmstadt.de

Phone +49 6151 16-20491
Fax +49 6151 16-29109

SFB 1053

DFG Deutsche Forschungsgemeinschaft

Viewpoint Feature Histogram (VFH) [1]



Summary

Future Work

Backup Slide

TECHNISCHE UNIVERSITÄT DARMSTADT

ESF