POSUDEK OPONENTA ZÁVĚREČNÉ PRÁCE

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	**Recognition of Road Traffic Participants in LiDAR Point Clouds**
Jméno autora:	**Čech Josef**
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Department of Cybernetics
Oponent práce:	Štěpán Obdržálek
Pracoviště oponenta práce:	Department of Cybernetics, Fakulta elektrotechnická, ČVUT

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání — **náročnější**

Hodnocení náročnosti zadání závěrečné práce.

The topic was well defined with good guidelines. The task was to prepare a suitable dataset of objects represented by sparse 3D point measurements, propose a suitable feature representation of the objects, and select and evaluate multiple of-the-shelf methods for classification of the objects.

Splnění zadání — **splněno s menšími výhradami**

Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.

The goals of the thesis were met in general. I have concerns about the data normalization method of choice, which in my opinion does not fulfill “The method should cope with […] variable rotation angle about the vertical axis” guideline. Also the request “The segments will be small point clouds of about tens to hundreds of points” was not really followed, the used data had an order of magnitude more points.

Zvolený postup řešení — **částečně vhodný**

Posuďte, zda student zvolil správný postup nebo metody řešení.

As mentioned above, my concern is about the chosen data normalization method, which, in my opinion, actually makes the classification problem harder. See notes in Section III.

Odborná úroveň — **A - výborně**

Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.

The topic required a good understanding of multiple topics from the computer vision field.

Formální a jazyková úroveň, rozsah práce — **A - výborně**

Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.

The thesis is written in good English, well organized and easy to follow.

Výběr zdrojů, korektnost citací — **A - výborně**

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

The sources are adequate.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uveďte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Overview:

The topic of the thesis is relevant and interesting. Classification of objects sparsely sampled by depth sensors is a hard task for several reasons. The objects are often not covered entirely, due to occlusions or limited vertical range of the sensors. Detection of the objects is hard and unreliable, as is the segmentation of the point cloud into point sets belonging to individual objects. The coverage density changes with object distance, and since only one side of the object is seen, the geometry of the captured points varies significantly with the object orientation w.r.t the view direction.

The classification task was simplified in the thesis by assuming well segmented points at the input. On the other hand, insensitivity to point density and to object orientation was called for in the thesis instructions. These are the most interesting and also the hardest questions of the assignment. The proposed feature representation raises some doubts about solving them well.

Concerns:

The goal of data normalization and feature selection is to minimize intra-class and maximize inter-class distances in the feature space. This increases class separation and simplifies the classification. My intuition is that the proposed data normalization actually does the opposite and reduces the class separation. This is somehow confirmed in the experiments, where the feature vector F_1, of which the author claims that it obviously does not separate the classes well (Figure 2.1), performs almost as well (or as bad) as the more rich feature vector F_2 computed on the normalized point clouds.

The chosen normalization seems to be designed for complete 360° object representation rather than for measurements taken from just one direction. Going step by step,

„the point cloud segment is shifted so that the center of mass is at the origin“ – that would work if it was a center of mass of the object. Center of the point cloud is view-dependent and will likely lay on, or close to, the surface of the object, making the representation view-dependent rather than class-dependent.

„the object is rotated so that dominant eigenvector of the covariance matrix is parallel to the x-axis“ - this is likely taking a vector almost parallel to the view direction, independently of the object class. Whether it is horizontal or vertical depends on the object, e.g. most trucks are taller than wide in frontal and rear views, so this operation would lay them flat on a side. Changing the view by perhaps 10°, the dominant eigenvector would change to horizontal, resulting in a significantly different representation of the same object.

„the segment is rescaled by a factor that is equal to the distance of the farthest point from the origin" – maximum being one of the least robust statistics, this makes the normalization very sensitive to perfect object segmentation. A single incorrectly segmented point could change the representation drastically.

It may be interesting to see how the classification performs without the normalization, especially the orientation part.

Questions:

The evaluation is rather coarse, coalescing the performance on a whole dataset to a few numbers. While that is OK for the purpose of comparison of different classifiers, with respect to the specifics of the point cloud data it would be interesting to see the results broken down to show some interesting aspects.

How does the performance change with the object distance/density of points? Does the classification still work when only "tens of points" are available, as asked in the thesis instructions? At what distance could objects still be classified?

What is actually the distance/density range in the datasets? Looking at Table 3.3 it seems that it is quite narrow. One could almost have a classifier based solely on the number of points on an object.

How is the classification sensitive to occlusions, or to only partially sampled objects?

Are there any object orientations where the classification fails? Figures in the thesis only show vehicles from angles where two sides are visible. Are all data like that (the main dataset is not public)? The classification is probably a much harder task when only rear or front sides of vehicles are observed.

How is the classification sensitive to the accuracy of the segmentation process?

It would also be nice to see a more fine-grained feature splits. Which features are the most useful for the classification and which do not bear any significance?

Formal:

The presentation is good. My only comment concerns the results shown in section 4. When performance results are presented in multiple figures, it is customary to show everything with identical ranges on axes. It allows for a quick comparison at a glance.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře.**

Datum: 31.5.2019 Podpis: Štěpán Obdržálek