CTU CZECH TECHNICAL UNIVERSITY IN PRAGUE

REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	Robot navigation driven by scene semantics
Author's name:	Jan Jirman
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Computer Science
Thesis reviewer:	prof. Ing. Jan Faigl, Ph.D.
Reviewer's department:	Department of Computer Science

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment — challenging

Evaluation of thesis difficulty of assignment.

The assignment asks to implement and experimentally verify semantic segmentation combined with terrain geometry estimation with a priori map information in autonomous navigation of a mobile robot to follow a sequence of waypoints with the allowed a large deviation of the sequence following. The assignment does not specify any possibly available building blocks; hence, it might be implemented from a scratch. Practical deployment and validation of autonomous navigation might be demanding, therefore, the assignment can be considered challenging because many building blocks exists and are available.

Satisfaction of assignment — fulfilled with major objections

Assess that handed thesis meets assignment. Present points of assignment that fell short or were extended. Try to assess importance, impact or cause of each shortcoming.

The assignment asks for a solution that would combine geometry estimation and a priori map. From the description in the text, it is not clear what is considered to be terrain geometry estimation. Regarding the cost function usable for navigation, it might be terrain cost related to robot motion capabilities and geometry of the terrain. However, the major part of the thesis is addressing the so-called contour, which is a border between binary classes' road and terrain, albeit it is not clear if contour also includes borders between road and car as defined in Section 4.2.2 (Page 15). There is no single word "geometry" in the text except the assignment and reference to documentation [22]. Furthermore, the assignment asks for collecting a datasets. However, there is no reference to the collected dataset, despite an indirect reference to "our dataset." Therefore, the assignment is considered as fulfilled with major objections.

Method of conception — partially applicable

Assess that student has chosen correct approach or solution methods.

The assignment does not specify the sensor needed, but the student selected RGB image-based segmentation only, which almost disqualified any suitable terrain geometry estimation. Related work is half of the page, and it is insufficient to choose a proper methodology. Since most of the referenced "publications" are links to the source code repository, it seems an available implementation has driven the student. However, about four thousand papers with code dealing with semantic segmentation[1] including approaches directly benchmarking road/street scenarios, such as SegNet from 2015, exist, just to point out that the addressed topic has been studied for a while by the community. Further, the defined concept of the contour aims to keep the robot navigated within the road (specifically to keep right, which is not sufficiently justified for left-hand traffic). However, the student completely omitted reactive approaches, such as visual road following[2], which works online in real-time, and successfully navigates a robot fully autonomously.

The assignment specifies that the given path (as a sequence of GPS-defined waypoints) might be followed with a large deviation from it. However, the student does not discuss that nor provide a control policy that would account for that. The assignment suggests employing a path follower in the navigation, a standard

[1] https://paperswithcode.com/task/semantic-segmentation
[2] https://doi.org/10.1109/ECMR.2015.7324212

way to navigate a robot, but the student considers it for future work. Instead, he tried to close the real-time loop with demanding image processing, which can be considered methodologically risky.

The discussion of the object in the scene detection is not sufficiently justified; regarding the assignment, it seems to be out of the thesis's scope.

From the provided text of the thesis, the student focused on employing selected implementation of the semantic segmentation with a learned model and building ad hoc solutions on top of the segmentation, which showed insufficient for the addressed task. Following GPS-defined waypoints assumes to have appropriately aligned local coordinate frame of the robot. Besides, it is also related to the localization capability of the robot, which is unfortunately not described in the thesis at all.

The text does not provide a dedicated problem statement with the discussion of the made assumptions and robot setup. In Section 4.2.3, the student presents that the system is supposed to be deployed on flat terrain, because the used transformation assumes that. It makes the terrain geometry questionable, as a flat surface does not have complex terrain geometry.

Because of the flaws above, the applied methodology is not considered correct. However, based on the further discussion with the students, some unclear parts have been clarified. Therefore, the methodology is assessed as partially applicable.

Technical level — E - sufficient.

Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.

The student attended the artificial intelligence in robotics course, where autonomous navigation of the robot is structured into the determination of the waypoints, path planning, and path following, which also fits the addressed problem. Fig. 3.4 includes a path tracker, but for some reason, the student did not elaborate on it or describe the control policy adequately. Furthermore, the course includes an overview of the control architecture and control loop frequencies that should be considered in the proposed solution. The described issues with a 2 Hz image processing loop are not significant depending on the robot speed, which is unfortunately not mentioned. Here, one can expect an analysis of the field of view and processing time that would yield the maximal expected speed the system can manage. However, such an analysis is not provided in the thesis's text.

Further, having delayed measurements is the default setup of a robotic system operating various sensors and control loops running at different frequencies. It might not be a surprise. It is one of the very reasons to employ existing robotics middleware, where messages are time-stamped. The same holds for any localization pose estimate and the sensors' measurements integration into the world model. A transport delay, or latency, is a standard property of any robotics system. Therefore, there is no significant reason to struggle with that because standard methods exist to deal with it, such as waiting for the data. Even though it would be tedious, it will not affect the principle of the proposed method.

Having the knowledge of the available equipment and expertise in robotics research labs at the FEE CTU, some of the issues might be directly addressed. For example, using a total station for providing ground truth position and perfect initial robot alignment is relatively straightforward. Possible camera delays might be mitigated by using a dedicated camera. It is assumed that the camera inside the robot body has been utilized because the particularly used robot system is not described in sufficient detail.

The technical level is considered sufficient because the reported results provide sufficient evidence the developed solution have been deployed on a real robotic system.

Formal and language level, scope of thesis — D - satisfactory.

Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.

The text is hard to follow and contains many stylish issues that make the text almost unreadable. The text structure could be more suitable; it contains forward and backward references, making reading hard to focus. The abbreviations used need to be introduced appropriately. The addressed problem needs to be formally introduced and explicitly described. The first two chapters are single pages long. Chapter 3

describes available robots, but only one robot using an RGB sensor is used. The thesis lacks a formal problem statement with considered assumptions. Limitations are almost randomly mentioned, such as flat terrain, throughout the text making the puzzle complete. Equations are not numbered, which makes it hard to refer to them. The list of figures is three pages long because the additional short titles are not provided. Concerning the text in English by a non-native speaker, it can be considered satisfactory.

Selection of sources, citation correctness **E - sufficient.**

Present your opinion to student's activity when obtaining and using study materials for thesis creation. Characterize selection of sources. Assess that student used all relevant sources. Verify that all used elements are correctly distinguished from own results and thoughts. Assess that citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards.

The related work is very shallow, and most related approaches have probably not been considered. However, the selected approach showed to be sufficient to develop some solution and deploy it on the real robot. Most of the references include links to the repository with source codes, but none includes the date of the citations, making them not appropriately cited. The developed solutions are not included in the thesis but referenced in the repository, which is not publicly available. This reviewer has been able to access the code using his account only.

Additional commentary and evaluation

Present your opinion to achieved primary goals of thesis, e.g. level of theoretical results, level and functionality of technical or software conception, publication performance, experimental dexterity etc.

It is hard to imagine someone that can use the text of the thesis to replicate the results or continue on the initiated work. One of the most valuable findings during the thesis is the camera configuration to reduce the delay to about 0.3 s. However, the student considered it out of the thesis's scope, which is unfortunate.

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation. Please present apt questions which student should answer during defense.

Although the student made the system work, I feel uncomfortable assessing the text as satisfactory or better, albeit it might be relative. Many issues can be addressed differently regarding state-of-the-art or proper formal problem formulation. Hence, I am biased as the applied methodology did not yield satisfactory results regarding my interpretation of the assignment and stated expectations. I evaluate handed thesis with classification grade E - sufficient.

The presented thesis does not meet the expected standards of the master thesis. It needs to be adequately structured; it needs an explicit problem statement. The related work is shallow, and the selected methodology needs to be sufficiently justified. Regarding the state-of-the-art, the concept applied is questionable. However, the student implemented his solution and deployed it on a real robotic system, which provides sufficient evidence demonstrating that the student can address the given topic and develop a solution.

- The terrain geometry can be estimated using RGB-D or laser scanners. Why have these sensors not been considered, as they are available on the platforms listed in Chapter 3?
- What are the robot's speed and the minimal area covered by the field of view of the used camera to enable autonomous navigation using the proposed method?
- What is the proper configuration of the camera to reduce the delay?

Date: 6.6.2023 Signature