REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Indoor localization utilizing dense pointcloud**
Author's name:	**Anna Zderadičková**
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Computer Graphics and Interaction (K13139)
Thesis reviewer:	Torsten Sattler, Dr. rer. nat.
Reviewer's department:	Department of Cybernetics (K13133)

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **extraordinarily challenging**

Evaluation of thesis difficulty of assignment.

The thesis had two main goals: (1) generate a new dataset for visual localization from RGB-D data (captured by HoloLens devices), together with ground truth poses. (2) extend an existing visual localization pipeline by using depth data to fine-tune the initial poses estimated by the existing pipeline. On their own, each part by itself would have been a suitable thesis topic as each part requires a thorough review of the literature, knowledge from a rather broad range of research directions in the area of computer vision, is highly non-trivial to realize (when done properly), and requires detailed experiments. Combining both goals thus led to a highly challenging assignment.

Satisfaction of assignment **fulfilled**

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.

Even though the assignment is very challenging, the thesis succeeds in addressing all planned topics, including the two main goals.

Method of conception **correct**

Assess that student has chosen correct approach or solution methods.

The chosen method for generating the ground truth poses for the dataset covers all important parts of the alignment process (time synchronization, estimating the transformation between the two coordinate systems, estimating the poses of the markers in the coordinate system of the HoloLens, handling drift / accumulated errors in the poses). The proposed approach is suitable and technically sound. It certainly solves the problem.
The adaption of the HLoc localization pipeline, where different depth-based alignment methods are evaluated as post-processing, is technically correct.
The thesis follows standard evaluation protocols for the evaluation of the adapted localization pipeline on the developed benchmark dataset.

Technical level **B - very good.**

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

The assignment required that the student became familiar with a wide range of techniques (visual localization in general, pose refinement via 3D-3D correspondences, aligning images against detailed 3D models, evaluating camera pose accuracy, to name the most important ones) in order to tackle the posed challenges. It is clear from the thesis that the student succeeded in becoming familiar with these topics and was able to choose appropriate techniques.

Formal and language level, scope of thesis **B - very good.**

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

The thesis is well-written and easy to read. It follows a clear and logical structure. Equations are properly explained, and the individual terms are (mostly) clearly described. What I am missing in Chapter 4 is some motivation for why the proposed approach to generate the ground truth was chosen (and not other approaches (see below) that have been used

for this task), but that does not take away from the fact that the proposed approach is suitable for solving the problem at hand.
There is a major error in one of the measures used for evaluation (measuring the orientation error in Eq. 6.1). However, it seems likely that this is a mistake in writing the proper equation down as implementing the error measure in this way should lead to compile time errors. As such, I doubt that this error invalidates the evaluation.

Selection of sources, citation correctness **D - satisfactory.**

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.

This is the main weakness of the thesis. Contrary to what is claimed there are multiple (quite popular) datasets for indoor RGB-D navigation: 7 Scenes (2013), 12 Scenes (2016), RIO10 (2020). Unfortunately, neither of them is discussed in the thesis.
Generating ground truth poses for visual localization (and other tasks) is an established problem. Unfortunately, the thesis does not motivate the approach from Chapter 4 by contrasting it to prior work, e.g., the alignment procedures in [14], [Zhang et al., Reference pose generation for long-term visual localization via learned features and view synthesis, IJCV 2021], [Schops et al., A multi-view stereo benchmark with high-resolution images and multi-camera videos, CVPR 2017], or [Schops et al., BAD SLAM: Bundle Adjusted Direct RGB-D SLAM, CVPR 2019].
In general, the related work chapter often only explains what prior work has done, but does not make clear how it relates to the topics of the thesis.
The thesis clearly distinguishes between its own results and prior art.
Some of the citations point to arXiv versions of a paper rather than the peer-reviewed version published at conferences or In journals (e.g., [14]).

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.

I had a hard time following some of the experimental results in Chapter 4.1 as it is unclear to me how the curves are obtained from just a few measurements.
I found many figures in the thesis hard to read as they use a small font size and very thin lines.
I do not like citing a prior work as "Paper [XY] …". I would rather prefer naming the authors.

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.

The thesis addresses very challenging topics that is important for the computer vision community (generating benchmarks for visual localization, improving visual localization). In my opinion, it would have been better to split the thesis into two theses, with one focusing on the ground truth pose generation (and providing a more detailed analysis of its accuracy) and the other making contributions to novel localization algorithms. Still, I am impressed that the student was able to work on both topics in the presented level of detail in the given time. For a Bachelor thesis, the thesis covers a quite large and impressive range of topics that the student had to master. My main point of criticism is that the thesis fails to cite and discuss relevant prior work, in particular existing benchmarks and algorithms for aligning images to 3D models. Still, this does not take away that the thesis successfully addresses a

very challenging task by developing sound techniques for ground truth pose evaluation and by evaluating a wide range of pose refinement approaches on the proposed dataset. In my opinion, this is a very good thesis.

Questions for the defense:

- How does the proposed dataset relate to the 7 Scenes, 12 Scenes, and RIO10 datasets? What are its advantages and disadvantages compared to these datasets?
- Why not place the Vicon markers on the HoloLens farther apart to reduce potential wrong associations between the markers?
- Eq. 6.1 is incorrect. What is a correct way to measure the rotation error and was it used in the experimental evaluation?
- Why not use a more modern RANSAC variant (e.g., MAGSAC or Graph Cut Ransac) for the pose refinement process? Do you expect that more modern versions will have a major impact on performance?

I evaluate handed thesis with classification grade **B - very good.**

Date: **23.1.2023** Signature: