

## I. IDENTIFICATION DATA

<b>Thesis title:</b>	<b>Efficient Homotopy Continuation for Multi-View Geometry</b>
<b>Author's name:</b>	<b>Petr Hrubý</b>
<b>Type of thesis :</b>	master
<b>Faculty/Institute:</b>	Faculty of Electrical Engineering (FEE)
<b>Department:</b>	Department of Cybernetics
<b>Thesis reviewer:</b>	RNDr. Zuzana Kúkelová PhD.
<b>Reviewer's department:</b>	Department of Cybernetic

## II. EVALUATION OF INDIVIDUAL CRITERIA

<b>Assignment</b>	<b>challenging</b>
<i>How demanding was the assigned project?</i>	
<p>The topic of Petr Hrubý's master thesis is the application of homotopy continuation methods to efficient multi-view geometry computations. This is an interesting problem since many problems in computer vision, robotics, and other fields require solving complex systems of polynomial equations that are not efficiently solvable using algebraic methods, e.g., Gröbner bases or resultants. Using a modified version of homotopy continuation methods tuned for problems where the structure of the system is known in advance may improve the efficiency of existing solutions to many computer vision/multi-view geometry problems or may lead to efficient solutions for unsolved problems. Homotopy continuation methods are well studied in mathematics but have only recently been used to study or solve some more complex multi-view geometry problems in computer vision. These methods were either used as an offline tool to study the structure of problems, e.g., the number of solutions, or to solve the problem. However, existing applications of homotopy continuation methods do not use all properties of multi-view geometry problems, e.g., the fact that we are usually interested only in one geometrically feasible solution, we know the structure of the system that we are going to solve, we can generate an „infinite“ number of instances of the problem/system together with the ground truth solutions, etc. The goal of the thesis is to use such properties and suggest modifications of the standard homotopy continuation methods to achieve an efficient solution for problems in multi-view geometry.</p>	

<b>Fulfilment of assignment</b>	<b>fulfilled</b>
<i>How well does the thesis fulfil the assigned task? Have the primary goals been achieved? Which assigned tasks have been incompletely covered, and which parts of the thesis are overextended? Justify your answer.</i>	
<p>The thesis fulfills all three given tasks. The author applied a homotopy continuation method to two multi-view geometry problems, i.e., the 5-point relative pose and the four-points-in-three-views problems. The author implemented multiple modifications of the standard homotopy continuation method to achieve an efficient solution for the considered problems. In particular, tracking only one solution, selecting a starting problem-solution pair using a Neural Network Classifier, using Real Homotopy Continuation instead of Complex Homotopy Continuation, and an efficient evaluation of the linear systems contained in the predictor and corrector. This resulted in solvers running in tens of microseconds that are suitable for RANSAC. The solvers were tested on real datasets.</p>	

<b>Methodology</b>	<b>correct</b>
<i>Comment on the correctness of the approach and/or the solution methods.</i>	
<p>The thesis presents several contributions and demonstrates that the author has knowledge of homotopy continuation methods and multi-view geometry. Thanks to the suggested modifications, the proposed solvers are the first homotopy continuation-based solvers for multi-view geometry problems that achieve run-times of around 100 microseconds or less. Moreover, these modifications have the potential to be applied to multi-view geometry problems other than the two settings considered in this thesis.</p> <p>On the negative side, some of the used solution methods are not the most suitable and efficient ones and can be improved further. For example:</p>	

1. The proposed generators of problem-solution pairs are unnecessarily complex as they are running existing solvers, i.e., Nister's 5-point solver and P3P, to generate „ground truth“ relative poses while at the same time using ground truth provided by the datasets to select the correct solution among the multiple solutions returned by these solvers. Since the proposed method is interested in tracking only one solution, i.e., the geometrically correct solution, it is not necessary to run solvers that output all solutions. Such solvers might introduce some numerical instabilities, especially in close-to-degenerate configurations. Moreover, since there are no existing solvers for the four-points-in-three-views problem, the used combination of Nister's 5-point solver and P3P will, on noisy data, result in slightly different solutions than a proper minimal solver for this problem. As I understand, the only motivation for the proposed generators is to reflect the distribution of real data, including image noise. However, for generating training data for classifiers as well as anchors for the homotopy continuation method, I do not think that the modeling of image noise is necessary. It is sufficient to project the ground truth 3D points with ground truth poses from a real dataset to reflect the real distribution of point correspondences. On the other hand, the proposed generators of problem-solution pairs may be interesting for methods where more than one solution needs to be tracked.
2. The proposed invariantization of the problems and the alignment of the problem on the anchors is not the most efficient one. For example, the iterative approach for moving the center of mass to the origin can be replaced by an exact solution based on solving a system of polynomial equations. Similarly, the proposed approach for selecting a subset of the available permutations in the alignment seems ad-hoc. The main reason why this approach works seems to be the previous invariantization and the counterclockwise ordering rather than identifying the most successful permutations based on generating and evaluating lots of instances and tracks for the problem. The winning permutations preserve the first point (which was fixed on the y-axis) and the counterclockwise ordering of the remaining points. It seems a simpler solution than learning from data is sufficient here.
3. The use of a standard RANSAC implementation, as opposed to using a state-of-the-art RANSAC variant, with a fixed number of iterations seems to favor the proposed solver over Nister's 5-point solver. It is obvious that the proposed solver can at most achieve the accuracy of Nister's solver. The number of problems successfully solved by the proposed solver is only 22%. In contrast, Nister's solver, in general, solves nearly all problems (it is known that it has no problems with numerical instabilities). Therefore, RANSAC with Nister's solver will find a good model much faster than when using the proposed solver, which would allow RANSAC to terminate earlier. By fixing the number of iterations rather than using the standard adaptive termination criterion, the implementation is in favor of the proposed method in terms of run-times. Moreover, there are much more efficient implementations of Nister's solver than the one used in the thesis.

## Technical level

**B - very good.**

*Is the thesis technically sound? How well did the student employ expertise in the field of his/her field of study? Does the student explain clearly what he/she has done?*

The thesis is technically sound. The student demonstrated a good background in mathematics as well as multi-view geometry.

The main concern is that details on how homotopy continuation is used in the solvers are missing. For example, it is unclear whether the author implemented the full homotopy continuation method by himself or only plugged optimized solutions to the predictor and corrector step into an existing toolbox. Since homotopy continuation methods are not well-known to the geometric computer vision community, even providing instructions for using them that treat them as black boxes would be valuable.

## Formal and language level, scope of thesis

**C - good.**

*Are formalisms and notations used properly? Is the thesis organized in a logical way? Is the thesis sufficiently extensive? Is the thesis well-presented? Is the language clear and understandable? Is the English satisfactory?*

The presentation of the thesis can be significantly improved:

- There are many typos (I can provide detailed comments directly to the author).

- Since the four-point-in-three-views solver uses similar steps as the 5-point solver, Chapters 4 and 5 contain redundancies in the description, including (nearly) identical passages.
- A significant issue is that the notation used in the thesis is re-introduced in almost every subsection. Even though only one new symbol is typically introduced in a subsection, all previously used notation is re-introduced. The notation alone constitutes approximately 35 pages due to these repetitions. This makes the thesis much harder to read than it should be. It further makes it harder to concentrate on the main topics.
- Most of the notation is correct. However, there are some typos in the notation, and sometimes I would prefer slightly different indexing for better readability.
- Some parts of the thesis are unnecessarily detailed, e.g., Chapter 4.4, where even a simple thing such as the permutation of the points is described on two pages. Some other parts lack details, e.g., details about the classifiers and their implementation, as well as details on the homotopy continuation (see above).

The level of English is satisfactory.

### Selection of sources, citation correctness

**B - very good.**

*Does the thesis make adequate reference to earlier work on the topic? Was the selection of sources adequate? Is the student's original work clearly distinguished from earlier work in the field? Do the bibliographic citations meet the standards?*

The references are satisfactory. Even though the discussion of the state-of-the-art is short, it covers the most important works, including recently published papers.

### III. OVERALL EVALUATION, QUESTIONS FOR THE PRESENTATION AND DEFENSE OF THE THESIS, SUGGESTED GRADE

*Summarize your opinion on the thesis and explain your final grading. Pose questions that should be answered during the presentation and defense of the student's work.*

The thesis is a good submission, and it fulfills all its stated goals. The author demonstrates a good understanding of the fields of mathematics in the form of homotopy continuation methods and multi-view geometry. The proposed solvers, in my opinion, are not yet fully practical, i.e., the proposed 5-point solver is less stable and slower than Nister's and the 4-point solver is quite unstable. Still, the thesis is an important step towards practically relevant solvers based on homotopy continuation. In summary, the topic of the thesis is important to the field; the thesis goals were met, and interesting results were achieved. I recommend the thesis for defense and propose a grade of B (very good).

Additional comments and questions:

1. It is not clear how the classifier was evaluated. According to Algorithms 11 and 17, the training data contains multiple labels for the same problem representation. The classifier returns a vector of probabilities, and the anchor corresponding to the highest probability is selected. In the evaluation of the classifier, only 20% respectively less than 10% of the validation problems are correctly classified, which seems very low. It is not clear what is meant by "correctly classified" since one problem can have multiple labels, i.e., one problem can be tracked from multiple anchors. Could the low success rate of the classifier be due to the way multiple labels for the same problem are handled?

2. The percentage of the problems successfully solved by the proposed solvers is also quite low (22% for the 5-point problem and even less for the 4-point problem). It would be interesting to understand the reason for this behavior. There can be several reasons, e.g., tracking the wrong solution, tracking in real instead of complex numbers, or starting from a wrong anchor. Experiments that compare the proposed method with methods that are, e.g., tracking one solution but in complex numbers or tracking more than one solution, would be interesting.
3. The mean average accuracy measure used in the experiments (cf. Algorithm 18) does not seem to correspond to a mean average measure (as, for example, used in image retrieval). Rather, it simply seems to be the fraction of errors below 10 degrees.
4. The experimental evaluation is missing details and important experiments. For example, how are the solution accuracy and run-time correlated with the number of anchors? How does the proposed 4-point approach compare to a simple baseline combining 5-point relative pose with a P3P solver? Further, more sophisticated RANSAC experiments (see above) would provide a better insight into the practicality of the proposed solvers.
5. Why not take the length of the track and the quality of the solution into account when training a classifier to predict the most promising anchor(s)?
6. During the generation of the anchors (see Section 6.2.1), were the 10000 points (problem-solution pairs) generated randomly? It would be interesting to have problem-solution pairs that reasonably cover the space of all possible relative poses and point configurations encountered in the real world. This can be, e.g., achieved by combining images from different datasets that have different distributions of points and relative poses.
7. One interesting approach to increase the percentage of successfully tracked problems is to adaptively switch between multiple classifiers that use different numbers of anchors inside RANSAC. For example, one could start with a classifier using 10 anchors and switch to one using 100 anchors if the percentage of successfully tracked problems is too low.

The grade that I award for the thesis is **B - very good**.

Date: **9.6.2021**

Signature: