Report on the thesis

Camera Pose Estimation from Lines using Direct Linear Transformation

by

Ing. Bronislav Přibyl

The thesis studies absolute internally calibrated camera pose computation from images of points and lines. This is a classical problem of computer vision that attracted considerable attention. A large number of algorithms for computation of camera pose from either points or lines has appeared. Despite that, not many methods for using points and lines together have been known before. Using lines and points together is important for reducing the occurrences of critical configurations and for allowing for larger fraction of wrong matches. The goal of the thesis is to develop a new algorithm that would be able to use points as well as lines and would be efficient for a large number of measurements. A new algorithm based on "Direct Linear Transformation" (DLT) is presented and compared to the state of the art.

Chapter 1

Chapter 1 reviews general background and goals of the thesis.

Chapter 2

Chapter 2 provides presents basic concepts for (i) the geometry of perspective cameras, (ii) representation of points and lines in space, (ii) their projection into images, (iv) detection and matching of lines in images (v) and solving linear equations. In general, the description is presented correctly. However, the description of solving linear equations is somewhat too simplified and, at the same time, so standard that it would be better to refer to a standard literature.

Chapter 3

Chapter 3 first gives some motivation for using lines for camera pose computation, which are reasonable and well founded. Then, it provides camera pose estimation from lines. It first reviews the local optimization (iterative) methods. Then it deals with algebraic methods which are based on solving minimal or almost minimal problems, sometimes followed by a local optimization refinement and often in a combination with, or as a part of, a robust matching scheme. In general review is correct but I would still like to see more separation of the pose computation from robust estimation.

Regarding the "algebraic methods", I miss the review of

[R1] Y. Kuang, K. Åström. Pose Estimation with Unknown Focal Length using Points, Directions and Lines. ICCV 2013 which presents several important minimal problems for computation of partially calibrated camera (all except for the focal length is known) pose from points, lines and directions. This is a very close and relevant work, which is missing in the review and should be compared.

Question 1: Can you please provide a comparison of your method to the approach in the [R1] and, in particular, comment on a possibility of developing a minimal solver for points and lines in your fully calibrated (normalized) case?

Next, methods based on solving linear equations only, named as DLT methods. The review is reasonable but I am missing the motivation for dealing with linear formulations separately from other formulations. In particular solvers using a small number of lines are close to minimal solvers and may be better or worse in stability as well as in computation time as it was observed in

[R2] Z. Kukelova, J. Heller, M. Bujnak, A. Fitzgibbon, T. Pajdla. Efficient solution to the epipolar geometry for radially distorted cameras. ICCV 2015.

Question 2: What is so appealing on fully linear methods?

Finally, the chapter ends with a brief review of RANSAC and Reweighted Least Squares. It is a very short exposition of very familiar material that could better be placed in Chapter 2 or replaced by a reference to a standard literature.

Question 3: The last sentence on page 26 says that 60% breakdown point was claimed in [22]. Has this been ever verified by a synthetic experiment with realistic distribution of gross errors?

Chapter 4

Chapter 4 presents the new contribution of the thesis. It starts with another review of the state of some of the art methods that tries to evaluate their strengths and weaknesses. The review addresses important works in the field but I find it somewhat confusing. Current paradigm in geometric estimation in presence of gross errors in computer vision is based on

1. Using (almost) minimal problem estimators providing candidate solutions quickly and to be efficient in robust estimation based on RANSAC and its variations.
2. Local optimization embedded in RANSAC to cope with small samples and to reach a good initial estimate.
3. A few steps of full non-linear bundle adjustment to get a best possible local minimum from the initial estimate.

Hence it makes sense to compare methods for each step above independently as well as their complete combinations in all three steps above. I do not see it clearly in the review.

Next, some common elements of the DLT methods are presented and explained in yet greater detail. Algorithm 2 is familiar from, e.g., [30] and need not be introduced again.

The main contribution of the thesis is presented in Sections 4.4.-4.6. It presents an elegant, unifying, and interesting formulation for using line-line, point-line and point-point correspondences within a single linear estimation framework. On the other hand, it brings new challenges in data normalization to give algebraic errors a quasi-geometrical meaning. The thesis faces these challenges and apparently brings a new normalization schemes that work. I believe that this is the strongest contribution of the thesis.

Question 4: We see (Table 4.2) that the new proposed method DLT-Combined-Lines still needs many line-line and point-line correspondences (5 lines + 10 points, …, 9 lines + 3 points). How is it when point-point correspondences are used? What do we gain by using line-line correspondences in combination with point-point correspondences compared to extremely efficient P3P algorithm (and its variations)?

Chapter 5

Chapter 5 deals with experimental evaluation of the presented approach and comparison to the previous work. I have the following comments:

1. It compares the thesis contributions to a number of previous approaches but, unfortunately, misses [R1] in this comparison.
2. In general, I find the experimental validation somewhat confusing. Many methods are compared but it is hard for a reader to really see when are the proposed methods really winning over the previous methods.
    a. Synthetic experiments, Figure 5.1., suggest that the new DLT-Combined-Lines and DLT- Plücker-Line methods become better (and very similar) only when using more than 10 lines. However, it compares the methods at, in my opinion, unrealistic detection errors. I believe that evaluation at 2 and 10 pixels detection error in 640x480 pixel image size is not interesting. Assuming reasonable viewing angle 60 degrees for 640 pixels, we get 0.33 degrees per pixel. This is a huge error in photogrammetry. I would be interested in a comparison for detection error in the range from 0.3 to 2.0 pixels.
    b. Figure 5.2. compares run times. It shows that some methods, e.g. "Ansar", are incomparably slower. I believe that this is not so interesting comparison. "Ansar" method is in fact almost a minimal problem and should not be used for more than a very small number of measurements in RANSAC scheme to generate proposals which would then be verified and updated by a local optimization step. In fact, I believe that the methods proposed in the thesis are best suited for the local optimization step in RANSAC. They are not robust, they need quite many measurements, and they are fast. To gain an insight into where is the real benefit of the thesis results, we should see a careful comparison of the methods inside the RANSAC robust estimation.
    c. The above note extends to the analysis of mismatches (Section 5.3) too. To be fair, we would have to see all pose computation method be evaluated within the same robust estimation method. This is possible but hard and hence an interesting alternative would be to show a very good combination of existing and new methods in the full state of the art RANSAC estimations method, e.g., based on

       [R3] R. Raguram, O. Chum, M. Pollefeys, J. Matas, JM. Frahm. USAC: a universal framework for random sample consensus. TPAMI 2013.

3. Table 5.3. shows the experiments with real data. I find it quite difficult to understand.
    a. Reprojection errors seem to be very strange We see values ranging from 10^-9 to 10^8.

Question 5: What are the units of the reprojection errors and how is it possible to obtain 10^-9 errors in real images. I would expect to see the smallest errors in the range 10^-1 … 10^0. Please explain.

4. Figure 5.8 presents the results after bundle adjustment initialized by the studied methods. I seem to see that the best methods are ASPnL, LPnL_Bar_LS, LPnL_Bar_ENull. Again, we see reprojection error in the range of 10^-10, which is never achieved n real data.

   Question 6: I do not see any real support for the claim in Section 5.5. that the new methods beat the other methods. Real experiments seem to show otherwise (however confusing they are). Please explain what is the scenario in which the methods introduced in the thesis provide the largest improvement compared to the state of the art.

Conclusions

The contribution of the work is, in my opinion, in the new formulation for computing camera pose from line-line, line-point, and point-point correspondences. This is a new result. It was first published in [II] at a conference in 2015. A similar method [68] was published later in 2016. The method presented in the thesis brings more elegant formulation than [68] by using the Plücker description of lines. I also value that the thesis provides experimental comparison to [68] and tries to show when it is superior.

The results of the thesis were published at British Machine Vision Conference in 2015 and in Computer Vision and Image Understanding journal in 2017, which proves that the contribution of the thesis was recognized by the computer vision community.

To conclude, despite some of the above criticism, Bronislav Přibyl carried out his own research, contributed to the state of the art by an original result and presented an acceptable thesis.

My recommendation is to grant the candidate the award of Doctor of Philosophy.

February 5, 2018

Tomas Pajdla

Associate Professor
Czech Institute of Informatics, Robotics and Cybernetics
Czech Technical University in Prague
pajdla@cvut.cz