Winter school project: Non-rigid Structure from Motion using Second-order Cone Programming

Yongbo Chen* supervised by A.Prof. Shoudong Huang*

June 29, 2021

Problem description: Non-Rigid Structure-from-Motion (NRSfM) is the problem of finding the 3D shape of a deforming object given a set of monocular images. NRSfM has been long treated in the literature as separate (different) problems from rigid structure-from-motion. When the scene structure is rigid the problem is generally well defined and has been much studied, with rigidity at the heart of almost all vision-based 3D reconstruction theories and methods. The NRSfM problem is naturally under constrained because there can be many different deformations that produce the same images. The common constraints, which include low dimensionality, local rigidity, and isometry, have been exploited to limit the set of solutions. As an open problem, the existing approaches for the NRSfM problem tend to be inaccurate, ill-posed, and unrobust. Existing NRSfM methods can be grouped into two main categories, Low-rank based methods and Physical based methods, depending on how deformations are modeled. A method belonging to Physical based methods will be introduced. As an example, the following image Fig. 1 shows the inputs and outputs of this problem:

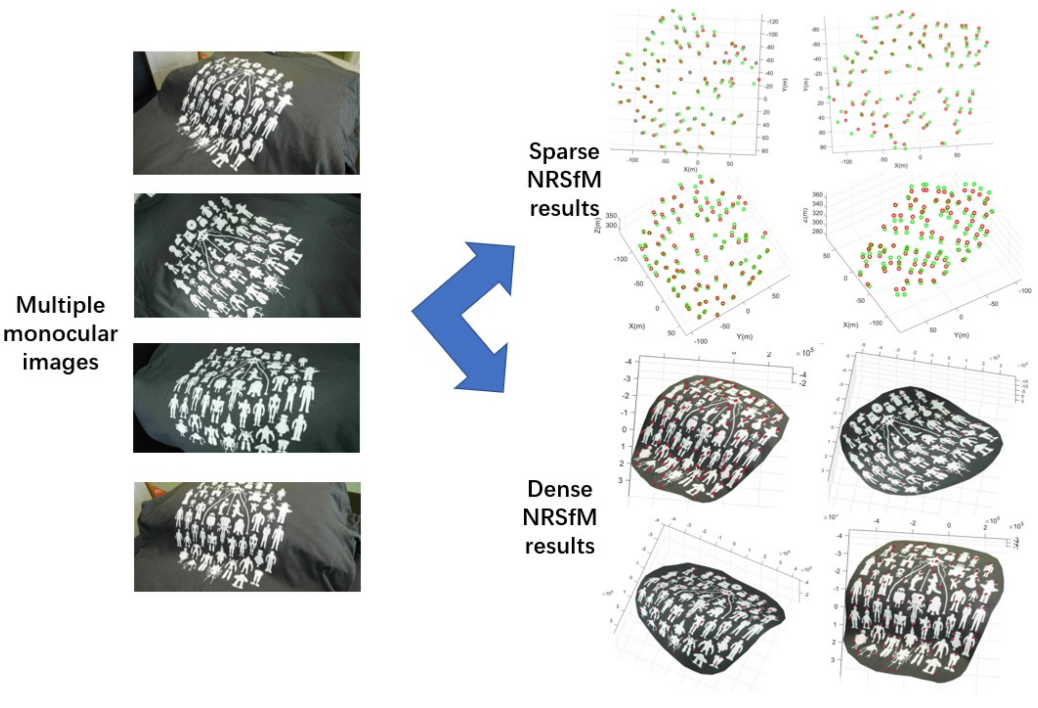

Figure 1: An example to show the inputs and outputs of NRSfM

Purpose: This project aims to introduce a Second-order Cone Programming (SOCP) based method, which is a convex formulation without requiring very good initialization, to solve the NRSfM problem. Based on this project, we would like to introduce the basic operations and solutions for this problem and further promote its application in the deformable SLAM.

Main reference: Paper: A. Chhatkuli, D. Pizarro, T. Collins and A. Bartoli. Inextensible Non-Rigid Structure-from-Motion by Second-Order Cone Programming[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 10, pp. 2428-2441, 1 Oct. 2018, doi: 10.1109/TPAMI.2017.2762669.

Code: Our deep learning tool is based on MATLAB and CVX toolbox. https://github.com/cyb1212/non-rigid-structure-from-motion-based-on-SOCP-Winter-school.git

Dataset: The datasets, including the flag, hulk, and T-shirt datasets, are shown in the code.

*Yongbo Chen and Shoudong Huang are with Faculty of Engineering and Information Technology, University of Technology, Sydney, Ultimo, NSW, 2007 Australia, e-mail: Yongbo.Chen@student.uts.edu.au, Shoudong.Huang@uts.edu.au.

Solution framework: The introduced framework is based on the inextensibility, which is a relaxation of isometry where one assumes that the Euclidean distances between points on the surface do not exceed their geodesic distances, and the SOCP convex formulation. The reconstructed 3D features are written as the depth as variables and 2D pixel feature. The NRSfM problem is solved by determining the unknown depth variables. Using a nearest-neighbourhood graph (NNG), the edges connected 3D features can be used to formulate the inextensibility constraints by a second-order cone form. In order to bound the problem, the scale of the intrinsic template is fixed by fixing the sum of the geodesic distances to an arbitrary positive scalar (1 in our case). Finally, we achieve robustness by introducing slack variables in the inextensibility constraint that can 'capture' outliers. The formulation changing are shown in the following image Fig. 2 and the readers can find their sepecific meaning in the main reference.

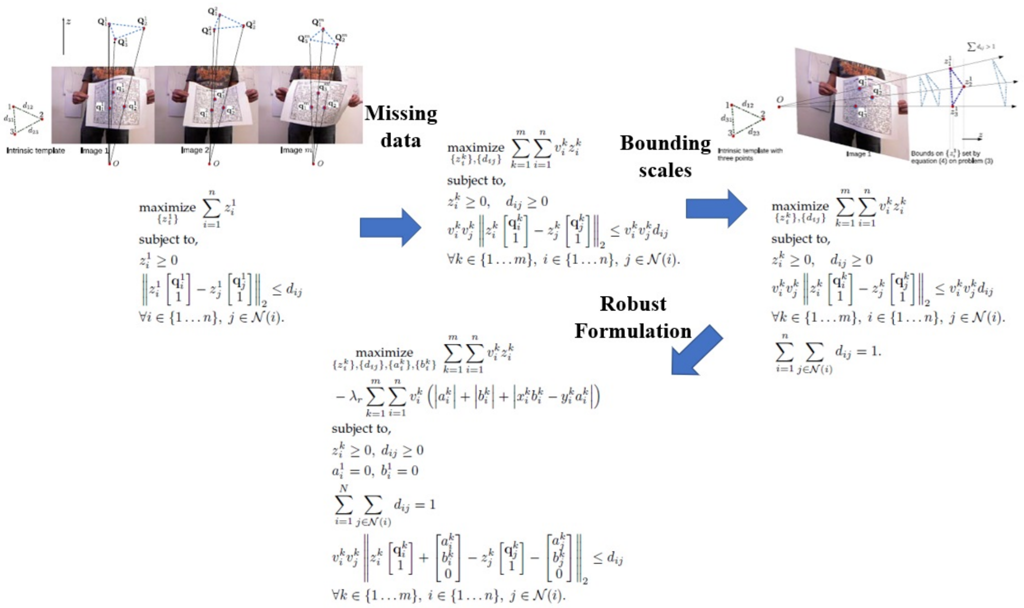

Figure 2: Model changing

Project step: The participants will be asked to complete the following exploration steps based on the provided code, including:

- **Adding constaints** In our provided code, the constraints, including scale limitation, second-order cone constraint, and positive limitation, have been deleted as a question for participants. You need to configure environment, adjust parameters, and fill this gap. The used codes need to follow the format of the CVX toolbox. This step is the basic and most difficult step in this project.

- **Testing more datasets** In our provided code, only the reduced T-shirt dataset with 4 scenes is offered. The participants are asked to add some other datasets in this project. The potential datasets include: (1) Flag dataset with 30 images, (2) full T-shirt dataset with 10 scenes; (3) Hulk dataset wiht 10 scenes.

- **Dealing with missing data** The case with some missing data is very common in the NRSfM problem and its later applications. The provided code is only suitable for the result with full dataset. If some of the feature correspondences are missing, the inextensibility constaints connected with them need to be delete. The readers is encourage to consider this sitation.

- **Testing new robust model** The basic problem formulation gives very good reconstructions when the input correspondences have no outliers. However in the presence of a few outlier correspondences, they break down easily. This is because the method does not model noise or errors in the point correspondences. We achieve robustness by introducing slack variables in the inextensibility constraint that can 'capture' outliers. The new formulation is presneted in the main reference.

Final evluation: The marks for these four steps are respectively: Adding constaints (60), Testing more datasets (20), Dealing with missing data (10), and Testing new robust model (10). The total marks are 100. Each part will be evaluated by combining the code and the final report.