Reconstruction Project

COL 780

November 2020

Using multiple views of a camera can enable the perception of depth. This allows us to estimate the camera position as well as the 3D location of points on a certain structure. In this project, we are going to use public sources of pictures (flickr, google images, bing images, etc.) of an Indian monument and reconstruct it in three dimensions by estimating camera poses and stitching together views from different cameras to infer the depth of corresponding points. There are a number of phases to the project:

1. **Data collection** - choose an Indian monument of your interest (except Taj Mahal since public datasets are available) and collect ∼500 images from public sources. Focus on high-resolution images.
2. **Generate a 3D reconstruction** of the monument using the default settings.
3. **Identify key weaknesses** of the reconstructions, e.g., lack of detail or arbitrarily scaled objects
4. Suggest **methods to correct** the above problems and validate them, building improved maps one after the other. Some of these techniques could be local/global bundle adjustment, choosing a different feature correspondence detector, changing the type of images acquired. Note that some problems could be corrected by modifying the input data itself.
5. **Submit a report** with all the modifications done and the corresponding changes with detailed evidence (in pictures) of the claimed improvements. Explain why the improvements are as expected, or in some cases, counter-intuitive.

Resources

1. You can start with the OpenSFM library[1] or the OpenMVG library[2].
2. Images can be collected from Flickr.
3. 3D vision tutorial[3].

4. Build Rome in a day[4]
5. Collection of recent papers on Multi-view geometry[5]

Details

1. The project has to be done in groups of two or individually.
2. The report should be written in a paper format. Use the template from [6].
3. The report should contain Introduction, Data Collection, Experiments (explicitly write problem, solution), Discussion
4. You are allowed to use publicly available code (both C++ and Python) with appropriate citation.
5. You are not allowed to discuss or borrow code from other groups.
6. Submissions are to be done on Moodle.

Evaluation Rubric

- Ability to run and explain code - 5 marks
- Data collection (and choices) - 5 marks
- Reconstruction improvement experiments - 15 marks
- Report - 5 marks
- **Bonus**: Collect pictures of your own house/nearby structure and use the same pipeline to show the improvement - 1 mark

References

[1] "Opensfm," https://www.opensfm.org/.

[2] "Open multi-view geometry github," https://github.com/openMVG/openMVG.

[3] "3d vision tutorial," https://github.com/sunglok/3dv_tutorial.

[4] S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S. M. Seitz, and R. Szeliski, "Building rome in a day," *Communications of the ACM*, vol. 54, no. 10, pp. 105–112, 2011.

[5] "Papers in multi-view geometry," https://github.com/openMVG/awesome_3DReconstruction_list.

[6] "CVPR 2021 template," http://cvpr2021.thecvf.com/sites/default/files/2020-09/cvpr2021AuthorKit_2.zip.