Fundamentals and Trends in Image Processing - IMPA
Professor and project's advisor: Luiz Velho
Student: César Morais Palomo (cpalomo[at]inf.puc-rio.br)

Project's theme: Community Photo Collections

Motivation

The use of community photo collections (CPCs), available through Internet photo sharing sites like Flickr and Google Images, has become a powerful new type of image dataset for computer vision and computer graphics research.

Whilst offering vast opportunities in research areas such as 3D reconstruction, image-based rendering, visualization and so many others, these CPCs also present researchers exciting challenges: these photos generally have extreme variability, having been taken by different photographers from myriad viewpoints, with varying lighting and appearance, and so many other intriguing characteristics for research purposes.

The challenge then becomes to seek algorithms that can operate robustly and successfully on such image sets to solve problems in computer vision and computer graphics.

Related work

The Community Photo Collections [1] group from University of Washington has been developing tools for processing and visualizing CPCs, usually aiming the average user and requiring no or little extra information but the input photos themselves.

As a breakthrough, 2006 Noah Snavely et al's *Photo Tourism* [2] was developed for browsing large collections of photographs in 3D. This tool automatically computes each photo's viewpoint and a sparse 3D model of the scene. Besides, a photo explorer allows the viewer to interactively move about the 3D space, through the photos, in a fashion and seamless way.

Further advances in the navigation control have been achieved by the same group, with the 2008 Snavely et al's paper *Finding Paths Through World Photos* [3] exposing details of how to discover a set of paths for traversing interesting regions and viewpoints of a scene, and how to turn them into controls for image-based rendering.

In a twist to the approach of navigating through the photos, but rather try to reconstruct the 3D geometry from these photo collections, Goesele et al's paper *Multi-View Stereo for Community Photo Collections* summarizes [4] how to take advantage of this kind of input data provided by the CPCs with remarkable results in 3D reconstruction.

Some work has also been done by the same group to try to address the problem of efficient structure from motion for CPCs, used in the mentioned works [5].

Proposal

The proposal of this project is to develop a full structure from motion framework to be used for CPCs, in a similar fashion to what has been done in *Photo Tourism* work.

Due to the limited time budget, the scope of the project must be limited, still trying to reach interesting deliverables by the end of the discipline's period. The brief of the planned project's steps are:

1. Features/keypoint detection in the input images using SIFT [6]
2. Features matching for each pair of images
3. Fundamental matrix estimation using the eight-point algorithm [7], using RANSAC [8]
4. Removal of matches that are outliers to the estimated fundamental matrix
5. Structure from motion step to estimate the parameters of each pair of cameras, in a bundle adjustment process

For each step, free libraries will be used to increase the chances of success.

At the end of the project, the intention is to have a basic SfM framework capable of registering the cameras used to take the input pictures, and also a basic visualization tool to play with the results.

References

[1] http://grail.cs.washington.edu/projects/cpc/
[2] Snavely et al - *Photo tourism: Exploring photo collections in 3D*.
ACM Transactions on Graphics (SIGGRAPH 2006 Proceedings), 25(3), 2006, 835-846.
[3] Snavely et al – *Finding Paths through the World's Photos*.
Proceedings of SIGGRAPH 2008.
[4] Goesele et al - *Multi-View Stereo for Community Photo Collections.*
Proceedings of ICCV 2007.
[5] Snavely et al - *Skeletal Graphs for Efficient Structure from Motion.*
Proceedings of CVPR 2008.
[6] David Lowe - *Object recognition from local scale-invariant features.*
Proceedings of ICCV 1999.
[7] Hartley and Zisserman – *Multiple View Geometry*.
Cambridge University Press, Cambridge, UK.
[8] Fischler and Bolles - *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography*.
Readings in computer vision: issues, problems, principles, and paradigms, 726-740 (1987).