Automated Panoramas from Multiple Photos of Multiple Subjects

Yuansi Chen Xie Lu Lingqi Yan

December 17, 2013

Abstract

In this project, we consider the problem of automatic construction of panoramas of multiple subjects. Previous approaches have used human input or restrictions on the image sequence for the matching step. Recent advanced approaches use object recognition techniques based on invariant local features to match images. These approaches provide a great improvement towards automatization of the task. However, they are computational inefficient and still require parameter tuning in model setting. We propose an algorithm based on local connection graph construction and spectral clustering, which completely automatizes the task with improvement in complexity.

1 Introduction

Panoramic image mosaicing has attracted great interest of research literature recent years. Especially with the development of commercial offers that come with digital camera, there is more and more demand on more intelligent and more automatic construction of panoramic images. In this project, we study the problem of automatic construction of panoramas of multiple subjects. It is common in life that we find multiple photos of multiple subjects in a folder after a touristic trip and we want to automatically assemble the pieces to make panoramas and remove outliers.

The geometry of the problem is well understood, and consists of estimating a 3×3 camera matrix or homography for each image. However, before getting into geometry, recognition is needed to identify connected images and classify images belonging to different subjects. Older approaches use human input or restrictions on the image sequence for image matching. It is painful for the user and restrictive to certain scenarios. Recent automatic approaches fall broadly into two camps: direct and feature based.

Lowe's paper [1, 2] based on local scale invariant feature (SIFT) [3] argued and demonstrated the advantage of feature based approaches when applied to a broader scenarios. Hence, our method is also going to be based on local scale invariant feature. However unlike these methods which use global feature matching scheme and probabilistic model verification,

which are computationally inefficient and require parameter tuning, we propose to improve these steps using a local feature matching scheme. We systematically compare our method to [1]'s approach in the rest of this report.
The rest of this report is organized as follows: Section 2 fixes the rigorous setting of the problem. In Section 3, we outline the overall framework and then proceed to describe in details. We evaluate our proposed method on a set of photos in Section 4.

2 Problem Setting

The input of the problem is an image set of n photos. These photos are supposed to be able to constitute K panoramas of K different subjects if ordered and assembled correctly. Noise photos which do not belong to any panoramas are also allowed. The following figure shows one input of our problem, where the set of 15 photos is unordered. These photos correspond to 4 panoramas with different subjects: Mount Aiguiller, Building, Sunset and Yosemite.

Figure 1: An image set of 15 photos. These photos are supposed to constitute 4 panoramas.

The goal of the problem is to automatically classify these photos by subject and then construct panoramas with them.

3 Automated panoramas of multiple subjects construction

3.1 Stage I: Feature extraction

The first step of our method is to compute SIFT features of all n images. Note that according to [3], SIFT features are partially invariant under affine change. This property greatly motivates the use the SIFT features in this problem as we know that correct match

consists of photos related by a homograph and small patches related by an affine transform (approximately).

3.2 Stage II: Feature matching and graph construction

Unlike the algorithm of [1], where a global feature matching (k nearest neighbours) as well as k-d tree construction is required to ensure that each feature is matched to its corresponding one in another image under the same subject, our method doesn't require global feature matching. There are two main reasons of avoiding global feature matching. First, the k-d tree should be computed each time a new set of images comes as input. Second, the construction of k-d tree requires $O(m \log(m))$ complexity where m is total number of SIFT features. m can be $1000 * n$ if we assume that there are approximately 1000 SIFT features per image. It can pose complexity problem when we scale up the number of photos.
For each pair of images, we find geometrically consistent feature matches using RANSAC to solve for the homography between them. Now for each pair of images, we know the number of consistent feature matches between them and how strongly they are connected. This number can be considered as the score of this candidate matching describing how probable two images belong to the same subject.
We obtain the weighted graph where each node is an image, and each edge is weighted by the number of consistent feature matches between its two nodes.

3.3 Stage III: Graph partition with spectral clustering

After obtaining the weighted graph, [1] attempts to delete weak edges, in order to get a graph where each connected components correspond to one subject, using a probabilistic model verification. It tries to discover a relation between the number of geometrically consistent feature matches (RANSAC inliers) and the number of non-consistent feature matches (RANSAC outliers) which provides an indicator of good image match or not. This method requires parameter tuning and is not reliable when the number of examples for parameter tuning is restricted. Our method, instead, leaves weak edges in the weighted graph. As we know that photos under the same subject are strongly connected in the weighted graph while photos from distinct subjects are weakly connected, we employ spectral clustering to automatically discover these strongly connected components.
We use the spectral clustering algorithm described in [4]:

1. Build the affinity matrix A directly from the weighted graph.
2. Define the diagonal matrix D and Laplacian matrix L, where $D_{ii} = \sum_{j=1}^{n} Aij$, and $L = D^{-1/2} A D^{-1/2}$.
3. Build matrix $U \in \mathbb{R}^{n \times K}$, whose columns consist of K leading eigenvectors $u_k (k = 1, ...n)$ of L.
4. Normalize the rows of U

5. Apply k-means to cluster the rows of U and assign the photo i to one subject accordingly.

Here is the heat map of the weighted graph before and after spectral clustering. We observe that spectral clustering does find the strongly connected components.

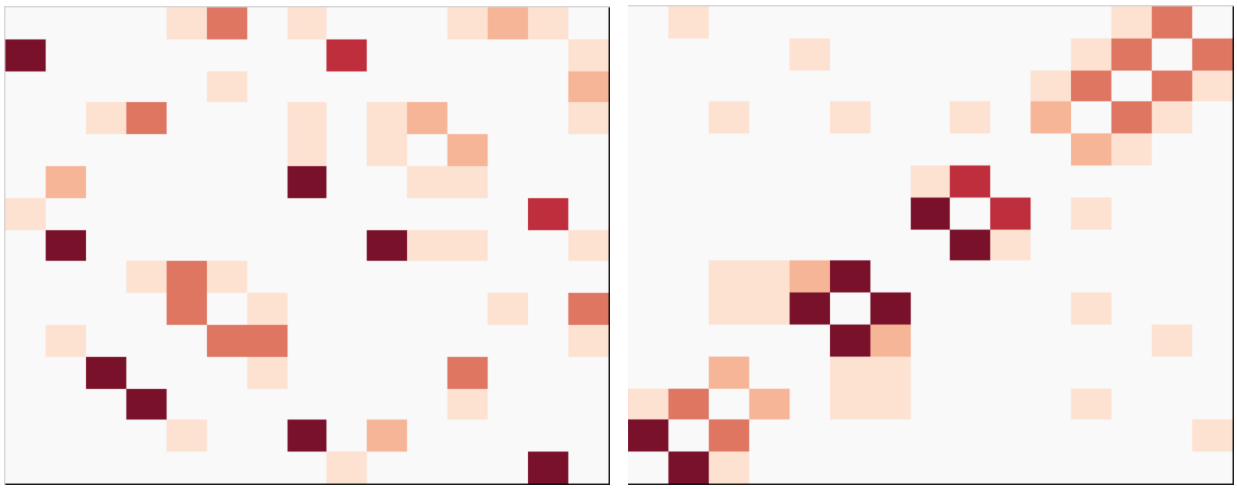

Figure 2: Heat map of the weighted graph before and after spectral clustering. The deeper the color, the stronger the connection is.

3.4 Stage IV: Homograph estimation and panoramas construction

Once photos are partitioned into K clusters and in each cluster all photos correspond to the same subject. We estimate the homograph two images by two images in the decreasing order of weights in the subgraph of the weighted graph. For estimating the homograph between two images, we use the eight-point algorithm described in class with RANSAC to compute the homograph and then verify the homograph by geometric verification.

3.5 Stage V: Blending

Ideally, each intersected patch would have the same intensity as the photos are taken at the same location and almost the same time. But in reality, it is not the case due to different lighting conditions and vignette effects. We use the simplest approach called "Feathering", or center weighting image blending. The pixel values in the blended regions are weighted average from the two overlapping images. Sometimes this simple approach doesn't work, for example, in the presence of exposure differences. But in our simple case, we are quite satisfied with the simple approach as this is not our main purpose of the project. More sophisticated approach such as that used in [2] could be considered in future work.

4 Results

The following figure shows the result of panoramas constructed from unordered photos. The next figure shows the result with blending.

Figure 3: Panoramas constructed from unordered photos.

Figure 4: Panoramas constructed from unordered photos. After blending, some edge lines are corrected.

References

[1] M. Brown and D. G. Lowe. Recognising panoramas. In *ICCV*, volume 3, page 1218, 2003.

[2] M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision*, 74(1):59–73, 2007.

[3] D. G. Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999.

[4] A. Y. Ng, M. I. Jordan, Y. Weiss, et al. On spectral clustering: Analysis and an algorithm. *Advances in neural information processing systems*, 2:849–856, 2002.