Panorama Stitching

ECE AIP Project 4, May 11, 2009

Wei Lu

1. Introduction

Panorama stitching is a technique that broadens the view of the captured scene by stitching correlated images with a proper warping model. In general, consumer camera only has a view of $50° \times 35°$, which is rather limited compared with that of the human vision ($176° \times 135°$) and a panorama mosaics (up to $360° \times 180°$).

In this project, we established a panorama stitching scheme with the help of various methods including feature extraction and matching by SIFT and outliner rejection by RANSAC [Brown and Lowe, 2007], linear model estimation with homography, and multiband blending with wavelets. The framework of our proposed scheme is shown in **Fig. 1**.

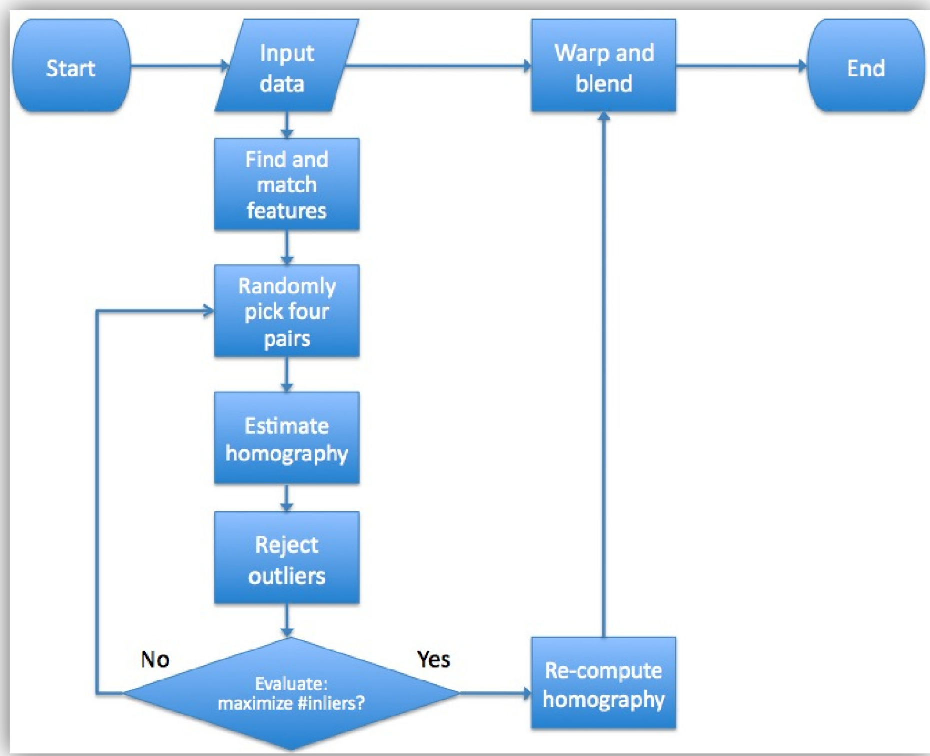


Fig. 1: the framework of the proposed panorama stitching scheme

2. SIFT

The Scale Invariant Feature Transform (SIFT) transforms an image into a large collection of local feature vectors, each of which is invariant to image translation, scaling, and rotation, and partially invariant to illumination changes and affine or 3D projection. It can be described with staged filtering approach:

- identifying key locations in scale space by looking for locations that are maxima or minima of a difference-of-Gaussian function;
- identifying collections of keys that agree on a potential model pose through a Hough transform hash table;
- fitting to a final estimate of model parameters through a least-squares approach.

A demonstration of SIFT features is shown in **Fig. 2**. Here, two corresponding images are connected seamlessly with each other. The green lines indicate the matching between points.

3. RANSAC

RANdom SAmple Consensus (RANSAC) is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers. The algorithm was first published by Fischler and Bolles in 1981.
The pseudo code for estimating homography by RANSAC is given in **Algorithm 1**:

Algorithm RANSAC
Do *#Trails* times
    Randomly pick four pairs of matched points;
    Compute the homography of the current trail;
    Collect inliers with $SSD < \varepsilon$;

    If the *#inliers* is larger than the incumbent record
        Update the information of the best set of inliers

Re-compute the homography using the best set of inliers
End Algorithm

Algorithm 1: RANSAC for estimating homography

A demonstration of feature point refinement is shown in **Fig. 3**. Similar to the way we present in **Fig. 2**, the green lines indicate the matching between points.

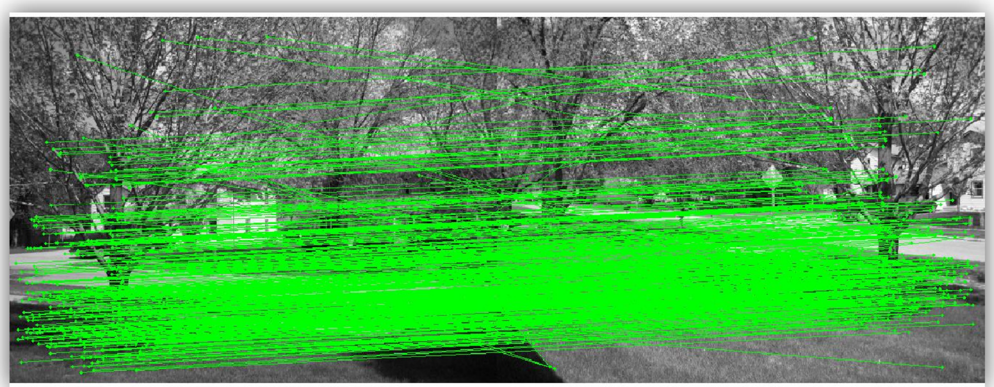


Fig. 2: SIFT features demonstration

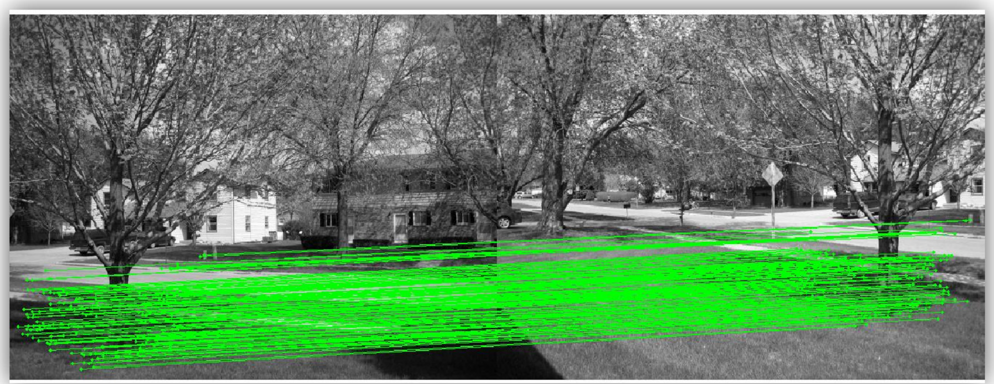


Fig. 3: outliers rejection by RANSAC

4. Perspective Warping

Perspective warping (a.k.a. homography) is a transformation combined with translation, rotation, scaling, and skewing (or shearing), etc. 2D homography can be expressed in a 3×3 matrix **H**. The warped point $\boldsymbol{x_2}$ is related with the original point $\boldsymbol{x_1}$ by

$$\alpha\, \boldsymbol{x_2} = \boldsymbol{H}\, \boldsymbol{x_1} \quad (1)$$

As the degree of freedom of **H** is eight, given at least four non-degraded pairs of matched points, we can estimate the homography from one image plane to the other one. Note the linear model assumes that the camera center is fixed with respect to all the projections; in another word, the result would be much more accurate if the scene is far from the camera. Our warping result is in **Fig. 4**.

(a) original image 1

(b) original image 2 to be warped

(c) warped image

Fig. 4: computed homography of two images

5. Multiband Blending

In our proposed system, images are blended in multiband. The main steps list as follows:

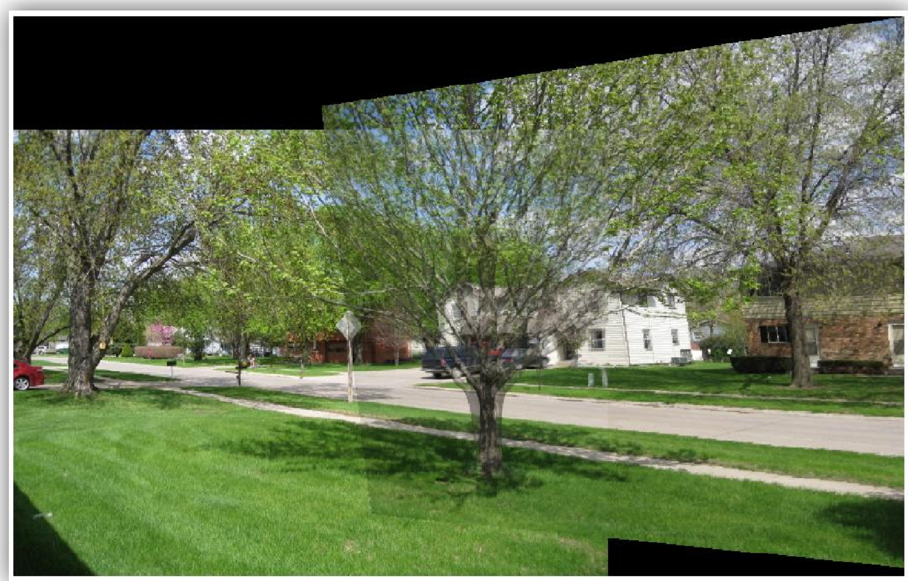
- decompose each image into different bands by wavelet transform
- compute the square distance map d from its center of each image
- compute the impact factor p of each image on a given point, where

$$p = \frac{d_2^2}{d_1^2 + d_2^2} \tag{2}$$

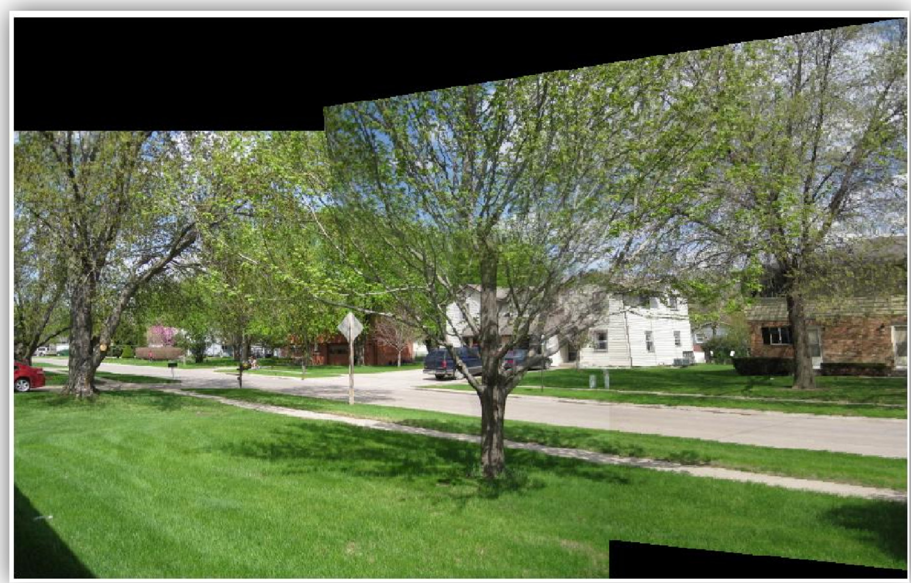
- blend corresponding components I_λ in the same band λ as a weighted combination, where

$$I_{b,\lambda} = p_\lambda I_{1,\lambda} + (1 - p_\lambda) I_{2,\lambda} \tag{3}$$

A comparison between direct blending in spatial domain and wavelet-based multiband blending is shown in **Fig. 5**. The corresponding wavelet decomposition (1-level with Haar filter) is shown in **Fig. 6**.



(a) spatial domain blending



(b) wavelet-based multiband blending

Fig. 5: comparison of two image blending methods

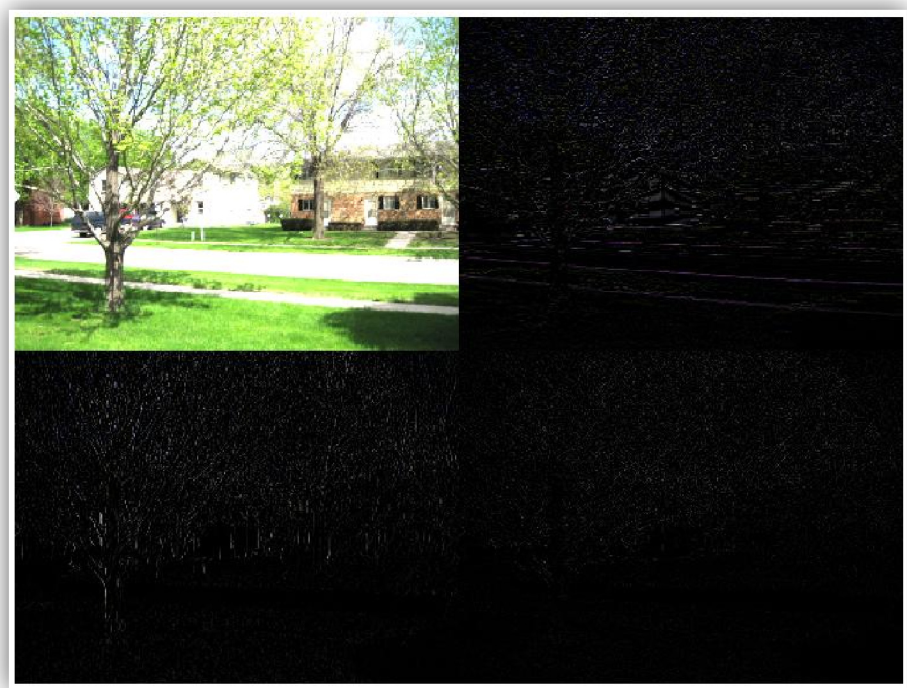


Fig. 6: 1-level wavelet decomposition (the original image is shown in **Fig. 4** (b))

6. Experiment Results

One example of stitching four correspondent images is shown in **Fig. 7**. The input images are given in (a); (b) and (c) are the stitching results by our proposed scheme and AutoStich [Brown and Lowe 2007], respectively.

(a) input images

(b) stitching by our proposed scheme

(c) stitching by AutoStitch

Fig. 7: panorama stitching results

7. Conclusions

The proposed panorama stitching scheme can stitch correspondent images correctly and smoothly. Further improvement can be made if we consider both white balancing and bundle adjustment.

8. Reference

Brown and Lowe, 2007. Automatic Panoramic Image Stitching using Invariant Features. *International Journal of Computer Vision.*