CS172 Computer Vision I: Homework 3 Report

Boyi Li
2018533187
liby@shanghaitech.edu.cn

Abstract

This LaTeX file is a homework report about image stitching. The first section is about the mathematical background about the method. It proves the correctness of the method. The second section is the result of the method. And the final is the conclusion.

1. Introduction

Mathmetics is no doubt the basis of modern digital image processing technique, some outstanding and popular image processing methods make use of the achievement of classic mathematical theorem. Image stitching is a technology that uses real images to form a panoramic space. It stitches multiple images into a large-scale image or a 360-degree panorama. Image stitching technology involves computer vision, computer graphics, digital image processing, and some mathematical tools. And other technologies. The basic steps of image stitching mainly include the following aspects: camera calibration, sensor image distortion correction, image projection transformation, matching point selection, panoramic image stitching (fusion), and brightness and color balance processing. Today we care about a simple image stitiching method using SIFT.

2. Background theorey

Our method mainly has 4 steps:

1) Feature Extraction
2) Image Registration
3) Compute matrix H using RANSIC
4) ImageWraping and Blending

2.1. Feature Extraction

It detects feature points in all input images. Image Registration Image Registration: establishes the geometric correspondence between images so that they can be transformed, compared and analyzed in a common frame of reference. Can be roughly divided into the following categories

1) Algorithms that directly use the pixel values of the image, for example, correlation methods
2) Algorithms that are processed in the frequency domain, for example, based on the Fast Fourier Transform (FFT-based) method;
3) Algorithms for low level features, low level features, usually use edges and corners, for example, feature-based methods,
4) Algorithms for high-level features, usually using overlapping parts of image objects, feature relationships, for example, Graph-theoretic methods

We use the SIFT method.

The SIFT algorithm is a scale-invariant feature point detection algorithm, which can be used to identify similar targets in other images. The image features of SIFT are expressed as key-point-descriptors. When checking image matching, two sets of keypoint descriptors are provided as input to the nearest neighbor search , and a closely matched keypoint descriptor is generated.

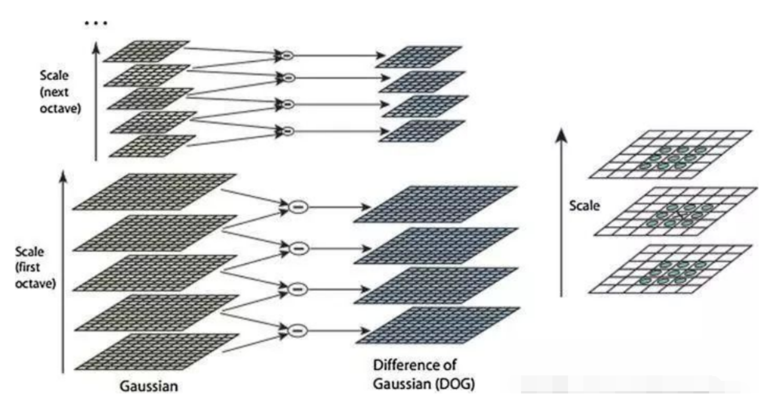

Figure 1.

The calculation of SIFT is divided into four stages:

1) Scale-space construction

2) Scale-space extrema detection

3) key-point localization

4) Orientation assignment and defining key-point descriptors

The first stage identifies potential points of interest. It uses the difference of the Gaussian function to search all scales and image positions. The location and scale of all points of interest found in the first stage are determined. The key points are selected according to the stability of the key points. A stable key point can resist image distortion. In the direction assignment link, the SIFT algorithm calculates the direction of the gradient around the stable key point. According to the local image gradient direction, one or more directions are assigned to each key point. For a set of input frames, SIFT extracts features. Image matching uses Best Bin First (BBF) algorithm to estimate the initial matching points between input frames. In order to remove unnecessary corners that do not belong to the overlapping area, the RANSAC algorithm is used. It removes false matches in image pairs. By defining the size, length and width of the frame, the reprojection of the frame is realized. Finally, stitching is performed to obtain the final output stitched image. When stitching, check whether each pixel in each frame of the scene belongs to the distorted second frame. If so, the pixel is assigned the value of the corresponding pixel from the first frame. The SIFT algorithm has both rotation invariance and scale invariance. SIFT is very suitable for target detection in high-resolution images. It is a robust image comparison algorithm, albeit slower. The SIFT algorithm takes a lot of time to run because it takes more time to compare two images.

2.2. Image Registration

After the feature points are detected, we need to associate them in some way, and the corresponding relationship can be determined by NCC or SDD (Sum of Squared Difference) methods.

2.3. Compute matrix H using RANSIC

Homography estimation is the third step of image stitching. In the homography matrix estimation, unnecessary corners that do not belong to the overlapping area are deleted. Use RANSAC algorithm for homography. The RANSAC algorithm fits a mathematical model from the observation data set that may contain outliers, and is an iterative method of robust parameter estimation. The algorithm is uncertain because it only produces a reasonable result with a certain probability, and this probability increases as more iterations are performed. The RANSAC algorithm is used to fit the model in a robust manner when there is a large amount of available data for laymen. The RANSAC algorithm has many applications in computer vision.

Estimate H with the RANSAC method:

1) First detect the corners of the images on both sides

2) Apply variance normalized correlation between corner points, collect pairs with sufficiently high correlation to form a set of candidate matches.

3) Choose four points and calculate H

4) Choose a pair consistent with homography.

5) Repeat step 34 until enough point pairs satisfy H

6) Use all point pairs that meet the conditions and recalculate H through the formula

2.4. Image Wraping

1) First calculate the deformed image coordinate range of each input image to get the output image size. It is easy to determine the output image size by mapping the four corners of each source image and calculating the minimum and maximum values of the coordinates (x, y) . Finally, it is necessary to calculate the offset xoffset and the offset yoffset of the offset of the specified reference image origin relative to the output panorama.

2) The next step is to use the reverse deformation described above to map the pixels of each input image to the plane defined by the reference image, and perform point forward and reverse deformation respectively.

3. Result display

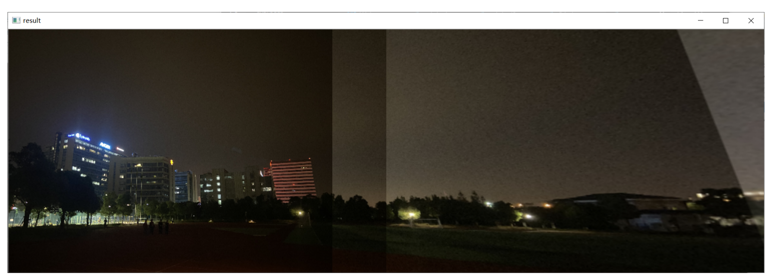

Figure 2.

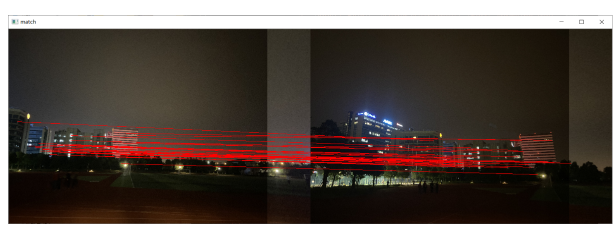

Figure 3.

Figure 4.

Figure 5.

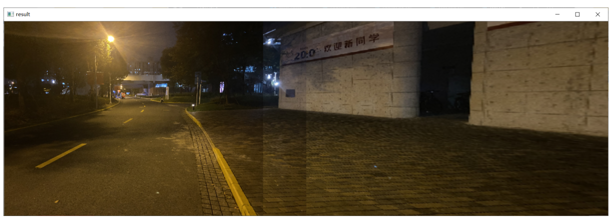

Figure 6.

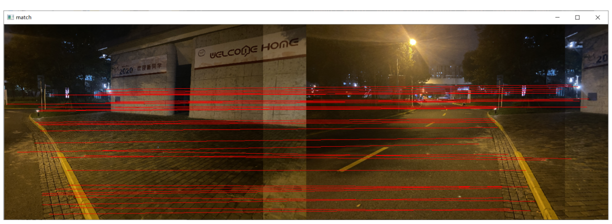

Figure 7.

References

[1] https://blog.csdn.net/u012384044/article/details/73162354