Automatic Image Stitching Using SIFT

Yanfang Li, Yaming Wang, Wenqing Huang, Zuoli Zhang
Zhejiang Science and Technology University, Model Recognition Library
E-mail:dami-li@163.com

Abstract

This paper concerns the problem of automatic image stitching which mainly applies to the image sequence even those including noise images. And it uses a method based on invariant features to realize fully automatic image stitching, in which it includes two main parts: image matching and image blending. As the noises images have large differences between the other images, when using SIFT features to realize correct and robust matching, it supplies a probabilistic model to verify the panorama image sequence. Addison to have a more satisfied panorama image, it uses a simple and fast blending method which is weighted average method. Finally, the experiment results confirm the feasibility of our methods.

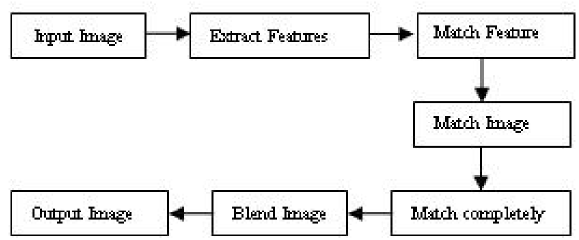
1. Introduction

Feature-based method [1, 2] is one of methods of Image stitching. And we propose a method based on invariant scale feature [3, 4, 5], which mainly includes two key parts: image matching and image blending. Image matching is used to find the motion relationship between two images or several images, and it directly relates to the success rate and the speed of the total process. While image blending is used to eliminate the various illumination of the adjacent image or color does not consecutive caused by the geometric correction or dynamic scene illumination. In that way two images can stitch into a seamless image.

In our paper, invariant scale features are also called SIFT features. SIFT features are local image features, which keep invariant in rotation, scale or illumination, and also robust in vision changes, affine changes or noises. Consequently, we use the method based on SIFT features, which solve the problem of image matching variant with scale variant in the method based on Harris corner detecting[6]. Otherwise, in the actual study, there are some noise images in the input image sequence. In order to solve this problem we develop a probabilistic model, by matching it to confirm the image and remove interference images. Furthermore, in order to eliminate the stitching visible seams and double edge of the panorama image we use the weighted average method for the image fusion.

2. The entire algorithm of the automatic stitching

The entire algorithm mainly includes: extract SIFT features; match features to get potential feature matches; match image sequence; match the image completely and blend the image, which can be described as the figure1.



Figre1.The entire process of this paper

2.1. Feature extracting

To each image, it builds image pyramid, and subtracts neighbor images to get the difference of Gauss (also called DOG) pyramid. Then, it detects the extreme for DOG pyramid. In order to locate the key points precisely to the sub-pixel accuracy, it fits quadratic function of 3D. At the same time, it eliminates the low contrast key point and the unsteady edge responses to improve the stability of the match. And the following step is orientation assignment, which uses orientation histogram to statistics the gradient orientation with sampling the center neighborhood of the key points. And the last step is to describe to the key points. Moreover, the detail process can be found in referees [3, 4].

2.2. Feature matching

We assume the camera rotates about its optical centre, and then the image would undergo a transformation, which can be written as:

$$\tilde{u}_i = H_{ij}\tilde{u}_j \qquad (1)$$

Where $\tilde{u}_j$ is the position before the transformation, $\tilde{u}_i$ is the position after the transformation, also, the parameter H_{ij} is deiced by rotational matrix and internal parameter of camera. The general form of H_{ij} is

$$H_{ij} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 0 & 0 & 1 \end{bmatrix} \qquad (2)$$

After extracting features, the next step is to match features by building a K-D tree to find k-nearest neighbors. In this paper, we find four nearest neighbors for each feature. In this way, we could get multiply feature matches.

2.3. Image matching

Based on multiply potential feature matches, we describe a modified RANSAC (random sample consensus) to match further. The modified RANSAC can be able to solve the motion relation between the potential feature matches efficiently and robustly, in which the points that satisfy the matrix are called RANSAC inliers, and others are called outliners. Moreover, using the connection of potential features and RANSAC inliers, we can build the probabilities model.

2.3.1. Modified RANSAC method

In our paper, it describes how to use modified middle filter to modify RANSAC. In that way, it can improve the effect of the RANSAC[7], from which we can get the matrix more robustly. As we known, middle filter is formed by two filters: angle filter and length filter. Consequently, we modify them to test the validity of the feature matching.

The rule of the modified angle filters are as following: selecting four feature matches randomly, and computing their entropy. If the entropy is larger than the threshold value, the feature matches can be selected to be the initialize feature matches, otherwise not. Furthermore, the entropy larger, the feature matches more easily picked to be the initial feature matches. The rule of the modified length filters is the same as the rule of angle filters. In shortly, the total algorithm can be shown as below:

(1) Selecting four feature matches, and using Modified angle filters to test, if satisfying the condition, going to the next step.
(2) Using modified length filters to test, if satisfying the condition, going to the next step. Otherwise, returning to the former step.
(3) Computing the transformational matrix.
(4) Testing whether the all possible feature matches satisfying the condition or not.
(5) Repeating the upper four steps for N times.
(6) Finding the feature matches that mostly satisfy the matrixes, and then computing the final matrix using them.

2.3.2. Probabilistic model

From the modified RANSAC, we get lots of features satisfying the affine matrix which are called RANSAC inliers. The method of our probabilistic model is to compare the probabilities of the inliers or outliners was generated by correct or false image match.

As below, we denote the event of the potential feature match i is inline or outline as f_i, and the image match is correct or not as m, which has only two values (0 or 1). Also we label the number of total features as n_f, and the number of inliers as n_i. According to the event f_i and $f_j (i \neq j)$ are independent, we can know that f_i obey to Bernoulli distribution. So the number of inliers of the overlap can be computed:

$$n_i = n_1 + n_0 \qquad (3)$$

where

$$\begin{aligned} n_1 &= n_f \cdot p(f_i \mid m=1) = n_f \cdot B(n_i, n_f, p_1) \\ &= n_f \cdot \frac{n_f!}{n_i!(n_f - n_i)!} p_1^x (1-p_1)^{n_f - n_i} \end{aligned} \qquad (4)$$

and

$$\begin{aligned} n_0 &= n_f \cdot p(f_i \mid m=0) = n_f \cdot B(n_i, n_f, p_0) \\ &= n_f \cdot \frac{n_f!}{n_i!(n_f - n_i)!} p_0^x (1-p_0)^{n_f - n_i} \end{aligned} \qquad (5)$$

In the above, p_1 is the probability of the event that the feature is inline under the correct image match; While p_0 is the probability of the event that the feature is inline under false image match.

According to the Bayes rule, we can calculate the condition that the number of inliers and total features should satisfy:

$$n_i > \alpha + \beta \cdot n_f \qquad (6)$$

If satisfy the above condition, the image match is correct, otherwise the image match is false. Moreover, the parameter α and β are decided by the parameter p_1 and p_0. In this way, we not only verify the matching relationship between images, but also reject the interfere image which is not belong to the panorama image.

2.4. Image matching completely

Above all, we have achieved the image match between two images. And in this stage we continue to research image match between multiply images. The key problem is to eliminate the accumulated error between the images, which direct to the success of correct image match. We use bundle adjustment [8, 9] to solve the above problem. Firstly, we choose one of the images to be reference surface. Then, each of other images transform to the reference surface, at the end of which all images are on the same surface.

The processes of bundle adjustment are as below: reading each of images into the adjustment, and constantly optimizing the parameters of the matrix in the adjustment. The method of optimizing is: Firstly, finding out the best neighbor image for each image, and directly calculating the distance between the two neighbor images. Then, mining the distance value to adjust the matrix between the neighbor images.

Given that there is one feature match (u_i, u_j) in the image I_i, the value of u_i which is projected to the reference surface and then to its neighbor image become to $\tilde{u}_i$. And the residual between them can be shown as:

$$r_{ij} = \left| u_j - \tilde{u}_i \right| = \left| u_j - H_j^{-1} \cdot H_i \cdot u_i \right| \qquad (7)$$

where H_i is the matrix transforming image I_i to the reference surface, and H_j^{-1} is the matrix transforming reference surface to the image I_j.

The object function is the sum of all the residual errors shown as equation (7):

$$e = \sum_{i=1}^{n} \sum_{j \in L(i)} \sum_{k \in F(i,j)} \left| f(r_{ij}^k) \right| \qquad (8)$$

where $F(i, j)$ is the set of feature matches between image I_i and image I_j. Then, we can update H_i:

$$H_i = H'_{ij} * H_i \qquad (9)$$

where H'_{ik} is the updating matrix value of image I_i and its best neighbor image. Furthermore, equitation (8) can be solved by L-M algorithm.

2.5. Image blending

The above stages realize image stitching in geometry, which has obvious seam in overlap region. It mainly due to ignoring the difference of illumining [10]. Our paper describes average weighed method which is simple and fast for blending, which can be described as figure2.

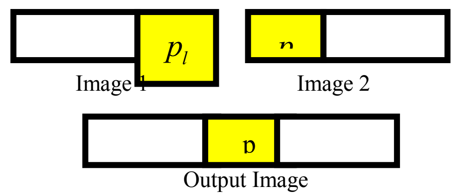


Figure2.Blending in the overlap region

In the average weighted blending, the values of features in overlap region are equal to the weighted average values of matching images, which can be shown as following:

$$p = \frac{d_l}{d_l + d_r} p_l + \frac{d_r}{d_l + d_r} p_r \qquad (10)$$

where d_l is the distance between the pixel in overlap region to the border of the left matching image, and d_r is the distance between the pixel in overlap to the border of the right matching image.

3. Results

Our paper uses one image sequence sampling in the park, which includes six images in size of 800×600, including a noise image, shown as figure3. And the result of automatic image stitching is shown as figure4. Seeing from the figures4, our method is well effective.

Figure3. One image sequence of the park

Figure4.The output of panorama image

4. Conclusions

Our method realizes automatic stitching of disorder image sequence which even includes noise images. Based on extracting invariant scale features, we get potential feature matches by k-nearest neighbor method, and then propose a modified RANSAC algorithm and a probabilistic model to realize image match precisely. Moreover, bundle adjustment is used to reach image match completely. And the average weighed method ensures smooth translation between the overlap regions.

While there are also some insufficiencies in our paper, such as noise would appear in low contrast color image stitching, and ghost would be created when there are moving object in the image sets. And these problems are needed to be solved in our future work.

5. Acknowledgement

This project is supported by National Nature Science Foundation of China with grant 60773024 and Zhejiang Provincial Natural Science Foundation of China with grant Y107558.

6. Reference

[1] C. Y. Chen and R. Klette Image stitching: comparisons and new techniques, Computer Analysis of Images and Patterns, 1999, 615-622.

[2] E. Vincent, R. Laganiere. An empirical study of some feature matching strategies. Vision Interface. 2002, 139.

[3] Lowe D. G. Object recognition from local scale-invariant features[C]. Institute of Electrical and Electronics Engineers Inc., 1999, 2: 1150-1157.

[4] Lowe D. G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[5] K.Mikolajezyk, C. Schmid. Scale& Affine invariant interest Point detectors. In IJCV, 2004,1(60):63-86.

[6] S. M. Smith, M. Brady SUSAN, A new approach to low level image processing. International Journal of Computer Vision.1997.23 (1):45-78.

[7] FISCHLER M. A, BOLLESR C. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6):381-395.

[8] P. F. McLauchlan , A. Jaenicke. Image mosaicing using sequential bundle adjustment. Image and Vision Computing, Vol. 20, Number 9-10, pp. 751-759, August 2002.

[9] Bill Trigg, Philip McLauchlan. Bundle Adjustment: A Modern Synthesis Vision Algorithms Workshop: Theory and Practice, pp. 298-372, 1999.

[10] T. Laakko, M. Mantyla. Introducing blending operations in feature models. Computer Graphics Forum, 12(3), p. C165-C176, 1993.