2010 Second Pacific-Asia Conference on Circuits, Communications and System (PACCS)

Study on Panoramic Image Stitching Algorithm*

Jubiao Li and Junping Du

Beijing Key Lab of Intelligent Telecommunication Software and Multimedia
School of Computer Science, Beijing University of Posts and Telecommunications
Beijing, 100876, China

ljb_233@163.com, junpingdu@126.com

Abstract **- This paper has a thorough study on panoramic image, an important application of virtual reality technology, and proposes a feature based method of generating a panorama. Descriptions of solutions and algorithms for each step were given in details, including Harris corner detection algorithm, RANSAC algorithm, Levenberg-Marquardt algorithm and so on. Some improvements of the algorithm were proposed in order to resolve the deficiencies. The experimental results were given at the end of the paper. It shows that the panoramic generated by this method is complete, clear, and no obvious distortion, and the methods adopted in the system are appropriate and practical.**

Index Terms – Virtual Reality; Panoramic Image; Image Stitching; Harris algorithm; Image Registration

I. Introduction

Panoramic image is an expression of virtual reality. It is evolved from images of real world, and generally has a full view and a high resolution in order to give users a lively experience [1]. Panoramic image is an important research direction and application of virtual reality technology. It involves digital image processing, computer graphics, computer vision, 3D engine, and many other technologies, and has a more and more widely future in virtual tourism, medical image analysis, digital city, etc.

One way to get a panorama is using some special photographic equipment, but these devices are often expensive and complex. Another way, which just needs an ordinary camera, is using a set of overlapping images to combine into an image that contains all the views. This method requires panorama stitching, which will be described below.

II. The Generation of Panoramic Image

The whole generation of panorama can be divided into three steps, see Figure 1.

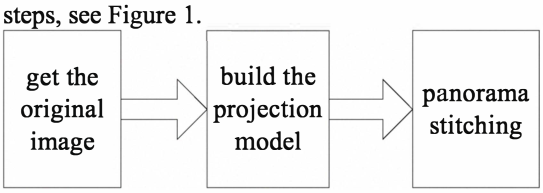

Figure 1 Panorama generation process

A. Original Image

Using an ordinary camera to get the picture is very simple. But there are two points that should be noted in order to make sure that the panorama will be made accurately. Firstly, every two images of the set should have some overlapped parts. Secondly, the horizontal high of the images should maintain a certain range.

B. Projection Model

Generally, panorama has three different projection models, cube, sphere and cylinder. Each of them has a different show result.

In this paper, we choose cylindrical model. It maps images to a suitable cylinder in order to get a 360 degree panorama. And the cylinder's core is the viewpoint. This model has two advantages. It is easy to shoot the original pictures and save the panorama, because the cylindrical surface can expand into a rectangular plane.

C. Image Stitching

The key of image stitching is image registration. The photorealistic and real-time ability all depend on the performance of image registration algorithm. According to the different way of image registration, image stitching technology can be broadly divided into three categories: phase correlation algorithm, region-based algorithm and feature-based algorithm.

Phase correlation algorithm is independent of scene, and can align translational 2D images accurately. Region-based algorithm chooses a window in the image as a template, and then contrasts the template with those images waiting for registration. However, this method is not suitable for the image which has a lager rotation or different perspectives. Feature-based algorithm selects a feature from these images, and then does the feature matching according to the principle of conformity. It has a good robustness to contrast degree and illumination change, also has a small computational complexity. How to select the feature is the key problem. And common features include Harris corner detection operator [2], SIFT characteristic, etc.

III. Stitching Algorithm

This paper studies on feature-based image stitching algorithm. The whole process is shown in figure 2.

* This work was supported in part by Beijing Natural Science Foundation of China (No.4082021) and Project of Beijing Education Committee.

978-1-4244-7968-9/10/$26.00 ©2010 IEEE PACCS 2010

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 12,2023 at 01:37:50 UTC from IEEE Xplore. Restrictions apply.

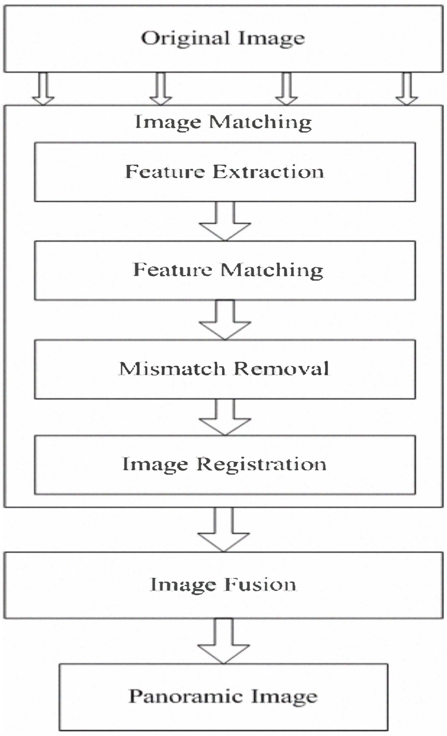

Figure 2 Stitching process

Mosaic of images is based on their overlapping regions, so the first problem is how to identify the overlap region. Feature-based matching algorithm extracts some feature points from the images, and then does the point matching. After removing mismatches, there is the image registration based on matching feature points. At last, those images, which have overlap region, inosculates to a panorama.

A. Feature Extraction

For the feature-based image stitching algorithm, how to select feature point has very important significance. The feature points often used to characterize the images need to satisfy two important conditions:

a. Feature points of the same scene in different perspective, viewpoint or illumination conditions, should be just the same.

b. The points should have a sufficient amount of information in order to match with each other.

Harris corner detector is an efficient feature detection operator, and has been used widely. It can quickly and accurately detect the corner point. Besides, it has rotation invariance and adequate information. However, the traditional Harris detector does not have scale invariance. That is, the same object with different sizes in different images will have different feature information. To solve this problem, this paper uses a multi-scale Harris corner detection algorithm.

1) Harris Corner Detection Algorithm

In Harris corner detection algorithm, grey change of regional image is described by first-order partial derivative. And it proposes the matrix M, which is associated with the image's autocorrelation function. And we can determine whether a point is a corner or not by calculating the eigenvalue of M. With the autocorrelation function to describe the sum of grey deviation of regional image, is as follows:

$$E(u,v)=\sum_{x,y} w(x,y)[f(x+u,y+v)-f(x,y)]^2 \quad (1)$$

W(x, y) is a gaussian filter. E's Taylor expansion is as followed:

$$E(u,v) \cong \begin{bmatrix} u\, v \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \quad (2)$$

In (2), M is a 2*2 symmetric matrix, as shown below.

$$M = \begin{pmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \\ \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) & \left(\frac{\partial I}{\partial y}\right)^2 \end{pmatrix} \quad (3)$$

I(x, y) is the grey value, $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$ are the image's first-order partial derivatives on x-axis and y-axis.

The curvature extremum value of image grey autocorrelation function on one point can be approximated by M's eigenvalue. If both of the two eigenvalue are large, that means curvature extremum value on x-axis and y-axis is large. In this situation, tiny movement will cause great changes in grey value, so that the point is a corner point.

The determinant of M ($\det M$) is in direct ratio with product of curvature extremum values, therefore, Harris corner detector can be expressed as:

$$R = \det M - k(traceM)^2 \quad (4)$$

$\det M = \lambda_1 * \lambda_2$, $traceM = \lambda_1 + \lambda_2$, λ_1 and λ_2 are eigenvalues of M, $0.04 \le k \le 0.06$.

When |R| is small, the point is in the smooth region of the image. When |R| is big, and R<0, the point is in the edge region. Only when |R| is big, and R>0, the point is corner point. Usually we choose a threshold T, when R>T, the point is a corner point.

Harris corner algorithm process is described as follows:

a. Calculate the first-order partial derivative on x-axis and y-axis of each point in the image. And then do Gaussian filter with its square and product.

b. Calculate the R value of each point. If R>T, we choose the point as a candidate feature point.

c. Among these candidate points, we choose the maximum in neighborhood as the final feature point.

2) Multi-scale Harris Detection Algorithm

Because Harris detector does not have scale invariance, we adopt multi-scale Harris corners [3] by using a Gaussian image pyramid. Specific algorithm is as follows:

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 12,2023 at 01:37:50 UTC from IEEE Xplore. Restrictions apply.

a. For each input image I(x, y) we form a Gaussian image pyramid $P_l(x, y)$ using a sub sampling rate s = 2 and pyramid smoothing width $\sigma_p = 1.0$

b. The Harris matrix at level 1 and position (x, y) is the smoothed outer product of the gradients. The output is:

$$M_l(x,y) = \nabla_{\sigma_d} P_l(x,y) \nabla_{\sigma_d} P_l(x,y)^T * g_{\sigma_i}(x,y) \quad (5)$$

The integration scale $\sigma_d = 1.0$ and the derivative scale $\sigma_i = 1.5$.

c. Compute the corner strength function:

$$f_{HM}(x,y) = \frac{\det M_l(x,y)}{trM_l(x,y)} = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} \quad (6)$$

In (6), λ_1 and λ_2 are eigenvalues of M, $f_{HM}(x,y)$ is also called the harmonic mean of the eigenvalues. Interest points are located where the corner strength $f_{HM}(x,y)$ is a local maximum in a 3*3 neighbourhood, and above the threshold T.

d. For each interest point, we also compute an orientation θ and an orientation vector:

$$[\cos\theta, \sin\theta] = u / |u| \quad (7)$$

U comes from the smoothed local gradient:

$$u_l(x,y) = \nabla_{\sigma_0} P_l(x,y) \quad (8)$$

The integration scale for orientation is $\sigma_0 = 4.5$. A large derivative scale is desirable so that the gradient field $u_l(x,y)$ varies smoothly across the image, making orientation estimation robust to errors in interest point location.

Since the computational cost of matching is super linear in the number of interest points, we can use adaptive non-maximal suppression (ANMS) strategy to select a fixed number of interest points from each image, and make sure that interest points are spatially well distributed over the image.

B. Feature Matching

The corresponding relation between two images' interest points can be found by feature matching. We can achieve this by comparing the feature point descriptor's error energy:

$$d = \sqrt{\sum_{i=0}^{127} (FV_1(i) - FV_2(i))^2} \quad (9)$$

We use the nearest neighbour algorithm, which is an improvement of k-D tree algorithm. Establish a k-D tree for each image, and make sure that the node of tree is the descriptor. Then search one image's descriptor in another image's k-D tree, till find all the pair of interest points.

C. Mismatch Removal

After the previous step, some pairs of interest points do not mapped to the same scene. This shows that they are not real matching points, but exterior points. The exterior point has bad influence on the calculation of parameter. So we need to remove exterior points. RANSAC (Random Sample Consensus) is the most widely used algorithm for removing the exterior points. By matching double points, the coordinate transformation relation of images can be work out. It is the transformation matrix M [4].

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (10)$$

$A' = M \circ A$, A and A′ are matching double points.

The steps of RANSAC algorithm are summarized below:

Select 4 pairs of matching points from N pairs randomly to establish equations. And work out the solution of M's unknown parameters. Then calculate the distance between the two points of N-4 pairs. If the distance is smaller than the threshold, the two points are interior points. Pick another 4 pairs, and repeat the process. The optimal solution is the parameter with maximum interior points.

D. Image Registration

The key problem of image registration is finding an appropriate model to describe the conversion relation between two images. Actually, we already have the transformation matrix M via RANSAC algorithm. However, projections of the same object in different images will be different since the projective direction is different. RANSAC algorithm does not do well in this problem [5]. Essentially, the problem of parameter optimization can be transformed into nonlinear least squares problem. In computational mathematics, a better solution of this problem is Levenberg-Marquardt algorithm [6].

1) The Cylindrical Projection

For the cylindrical projection, we find the corresponding projective point by solving parameter equations of plane image and cylinder. Refer to (11), w, h is the width and height of the plane image, f is the camera's focal length, (x′, y′) is the corresponding point on the cylinder.

$$\begin{cases} x' = f \times \arctan(\frac{x - w/2}{f}) + f \times \arctan(\frac{w}{2f}) \\ y' = \frac{f \times (y - h/2)}{\sqrt{(x - w/2)^2 + f^2}} + \frac{h}{2} \end{cases} \quad (11)$$

2) The Optimization of LMA Algorithm

When the error range of initial value was limited, the error of matching point will not be more than 0.1 pixels through the optimization of Levenberg-Marquardt algorithm. We take the error of interest point's coordinate position as the error function, and take the minimizing error of symmetrical projection as the optimization of matrix M. The error function is:

$$C = \sum_i d^2(x_i, M^{-1}x_i') + d^2(x_i', Mx_i) = \sum_i c_i \quad (12)$$

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 12,2023 at 01:37:50 UTC from IEEE Xplore. Restrictions apply.

Calculate the partial derivative of C at unknown parameter m. We obtain the Hessian matrix A and weighted gradient vector b.

$$a_{kl} = \sum \frac{\partial c}{\partial m_k} \frac{\partial c}{\partial m_l} \quad b_k = -\sum c \frac{\partial c}{\partial m_k} \tag{13}$$

Then add a correction value Δm to the estimated value of m. $\Delta m = (A + \lambda_i)^{-1} b$, and λ is a self-adjusting parameter. Iterate this process until C<T. And then we get a more accurate value of m.

In the experiment, we found that when the resolution was big enough, the algorithm has a higher success rate. When the resolution is lower, the convergence result may be wrong. Hence we proposed a comparative error of colour. When the resolution is small, if the error C decreases, but the colour error increase significantly, the direction of convergence is wrong. Choose the opposite direction in the next iteration. In order to avoid the negative influence of colour error, we require that only when C and total colour error reduced to a certain range, this rule can be available.

E. *Image Fusion*

We need to solve the problem of fuzziness and distortion caused by noise, rotation or offset in this step in order to satisfy the request of smoothness and seamless stitching [7]. We choose the slow in and slow out algorithm.

IV. Experiment Result

Experimental conditions: Windows XP system, 2.66GHz processor, 2G RAM, VC + +6.0

Figure 3 are four ordinary pictures, and figure 4 is the result. The results show that the panorama is complete, clear, and no obvious distortion.

Figure 3 Original image

Figure 4 Panoramic image

V. Conclusions

In this paper, we extract feature points by using multi-scale Harris corner detection algorithm, match by k-D tree, remove mismatch by RANSAC algorithm, and then, optimize the transformation matrix by the Levenberg-Marquardt algorithm, and complete the projection and image fusion. AT last, the experimental results show that the panoramic generated by this method is complete, clear, and no obvious distortion, and the methods adopted in the system are appropriate and practical.

Acknowledgment

This work was supported in part by Beijing Natural Science Foundation of China (No.4082021) and Project of Beijing Education Committee.

References

[1] Qingping Zhao. Summary of Virtual Reality. *Chinese Science*, January 2009.

[2] C. Harris; M.J. Stephens. A Combined Corner and Edge Detector. *In Alvey Vision Conference*, 1988. pp.147-152

[3] Matthew Brown; Richard Szeliski; Simon Winder. Multi-Image matching using Multi-Scale Oriented Patches. *IEEE,* 2005.

[4] Youqun Huang. Yu Fu. Guangkun Ma. Cylindrical Panoramic Mosaic Method Based on RANSAC Algorithm. *Journal of Shenyang University of Technology,* vol. 30, no. 4, 2008. pp.461-465

[5] Yu Tong. Dianxing Liu. Study on Several Related Issues of Panorama Algorithm. *Computing technology and automation,* vol. 25, no. 3, 2006. pp.111-114

[6] Min Yang. Study on Image Mosaic Based on Levenberg-Marquardt Algorithm. *Electronics engineer,* vol. 33, no. 10, 2007. pp.43-45

[7] Lifeng Jiang. Study on Key Technologies of Panorama Stitching. *Master's paper of Shandong university of technology*, April 2008. pp.56-57

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 12,2023 at 01:37:50 UTC from IEEE Xplore. Restrictions apply.