Proceedings of the 33rd Chinese Control Conference
July 28-30, 2014, Nanjing, China

Image Stitching Algorithm Research Based on OpenCV

CHEN Kaili[1], WANG Meiling[2]

1. Beijing Institute of Technology, Beijing 100081,China
E-mail: kelly_3224@163.com

2. Beijing Institute of Technology, Beijing 100081, China
E-mail: wangml@bit.edu.cn

Abstract: In view of the vast scene and high-resolution image stitching problems, this paper studies a kind of Harris image stitching algorithm which is based on OpenCV configuration environment. First of all, the Harris corner detection is used to extract the feature points. Then Normalized Cross Correlation is used to rough match the feature points and RANSAC algorithm is used to eliminate error matching. Secondly, cylindrical projection transformation model is adopted to realize image registration algorithm. And finally this paper uses an improved weighting average fusion algorithm to fuse images, reducing the computational complexity of image fusion and eliminating seam of stitched images. This algorithm reduces the overlapping rate of images during image stitching and increases the speed of image Registration and Mosaic. It realizes the purpose of giving consideration to both precision and efficiency.

Key Words: Harris Corner Detection, OpenCV, NCC, Weighted Fusion

1 Introduction

Image Stitching Technology is used in creating vast scene and high-resolution images which contain the image sequences information. It implemented registration and integration to a set of image sequences which mutual images have overlapping area. Image Stitching Technology has become a popular in the fields of computer vision and computer graphics. It has significant practical value of building a natural panoramic image.

Image stitching methods mainly include regional-based and feature-based. Region-based method[1] adopts relevant technology to match the corresponding position of a frame image pixels window in another frame image. This method is easy to implement, but it costs large amount of computation. Especially when the presence of rotating, zooming and translation transformation relation, blocking and degradation distortion images, the performance of algorithm would drop dramatically. Feature-based method[2] extracts common invariance features, such as contours, moments, and then accurate match. As the result of the features of invariance and unrelated to the gray-scale image, it can effectively solve the ambiguity of registration. Feature-based method due to its robustness has been widely used. Most prior research has focused mainly on extracting feature points as corners, such as Harris corner algorithm and SUSAN corner algorithm.

Firstly, Traditional Harris corner algorithm, which needs three Gaussian smoothing, a relatively large amount of computation, and sensitive to noise, has poor matching accuracy[3]. Currently the vast majority of splicing methods focus on improving the precision of image registration research, but less study on the efficiency of matching. Improve the matching accuracy can improve the quality of the panorama image, with larger than the calculation of image matching, speed stitching algorithm has not been improved.

This paper bases on the practical camera model and study the panoramic image mosaic model. On the premise of guarantee the quality of image, the paper focuses on the problem of the premise of ensuring image quality and the speed of image stitching. Specifically, this paper focuses on the following two aspects: first, improving the cylindrical projection process, which includes inverse projection and distortion correction using bilinear interpolation. Second, the paper designs a panoramic image stitching system prototype, for stitching panoramic image. The algorithm overcomes the limitations of transform domain-based approach to improve the accuracy of feature matching. For the existence of translation and rotation of the image at any angle are able to achieve a smooth mosaic.

2 Improved Cylindrical Projection

Image acquisition tool used in this paper is five closely spaced the same wide-angle camera on the pentagon, the image captured is 616 * 807. The camera has been calibrated, which the camera focal length f in the horizontal direction is 386 in pixels. h_{fov} is the center angle of each image, f is the camera focal length in the horizontal direction in pixel, W is the image width in pixel, H is the image height in pixel. By formula (2) the center angle h_{fov} of each image can be obtained. Since the camera is wide-angle with barrel distortion, the initial images have been correction well ahead of time.

$$f = 386 \tag{1}$$

$$h_{fov} = a\tan\left(\frac{W}{f}\right) \tag{2}$$

*This work is supported by National Natural Science Foundation (NNSF) of China under Grant 00000000.

2.1 Cylindrical projection

Cylindrical projection mode refers a sequence of collected images projection transform into cylinder that has fixed focal length radius, and the panoramic image stitching bases on this cylinder. This model has a $360°$ visual angle in the horizontal direction, and less than $180°$ rotation angle in the vertical direction.

Cylindrical projection has several advantages:

(1) Image acquisition mode with respect to the cube model and spherical model should be simple and does not require precise positioning.

(2) Cylindrical projection image can be expanded into a rectangular plane image, and the image format is convenient for computer accessing.

Cylindrical surface equation:

$$u^2 + v^2 = f^2 \tag{3}$$

Assuming the projection point A_1 coordinate of point A in cylindrical projection coordinates is (u, v, w), the expression of the parameters as follows:

$$t = \frac{1}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \tag{4}$$

$$u = \frac{f\left(x - \frac{W}{2}\right)}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \tag{5}$$

$$v = \frac{f\left(y - \frac{H}{2}\right)}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \tag{6}$$

$$w = \frac{f^2}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} \tag{7}$$

To obtain the cylindrical coordinate of images, it is need to translate three-dimensional coordinate to two-dimensional coordinate.

$$\bar{x} = f * \arctan\left(\frac{x - \frac{W}{2}}{f}\right) + f * \arctan\left(\frac{h_{fov}}{2}\right) \tag{8}$$

$$\bar{y} = \frac{f\left(y - \frac{H}{2}\right)}{\sqrt{\left(x - \frac{W}{2}\right)^2 + f^2}} + \frac{H}{2} \tag{9}$$

2.2 Improved Cylindrical projection

The texture of photos that captured by cameras is often distorted, which cannot be directly used for texture mapping. It is important to study the texture distortion correction. Inverse projection algorithm is presented to rectify distortion for cylindrical objects. This algorithm displays the image texture projection on the cylinder of the two-dimensional plane and takes advantage of linear interpolation algorithm to correct defect pixel values. The parameter expression of inverse projection points $A_3\left(\bar{\bar{x}}, \bar{\bar{y}}\right)$ is as follows:

$$\bar{\bar{x}} = f * \tan\left(\frac{\bar{x}}{f} - \frac{hfov}{2}\right) + \frac{W}{2} \tag{10}$$

$$\bar{\bar{y}} = \frac{\left(\bar{y} - \frac{H}{2}\right) * \sqrt{\left(\bar{x} - \frac{W}{2}\right)^2 + f^2}}{f} + \frac{H}{2} \tag{11}$$

The vertical coordinates of the point A and A_1 do not changed. Similarly, the scene vertical does not change after cylindrical coordinate transformation. After matching feature points and fusion image, it is need to transform mosaics of images to the original coordinate, namely cylinder inverse transformation. The above expressions can be applied in the acquisition coordinates projected to (u, v, w) parametric coordinate system. A complete panorama needs another four images projected onto $\left(\frac{4u}{5}, v, w\right)$、$\left(\frac{-3u}{5}, v, w\right)$、$\left(\frac{-2u}{5}, -v, w\right)$、$\left(\frac{u}{5}, -v, w\right)$ in turn to constitute the panorama of information. The center angle that exceeds $90°$ of each image is projected on two adjacent coordinate systems, as the overlap region for matching and integration.

3 Image Stitching

3.1 Harris corner Detection

Harris operator is C.Harris and M.Stephens in 1988 proposed a point feature extraction operator [4]. This operator is inspired by correlation function in signal processing, which provides matrix M that linked autocorrelation function. Eigenvalues of matrix M is a first order curvature of autocorrelation function. If the curvatures of the both Eigenvalues are high, then this point is considered as corner point. This method has good robustness for image rotation, brightness, viewpoint change and noise.

Operate the image at gray level and take a target pixel as the center of a small window. The window moves along four directions of up, down, left and right. Calculate the gray change in four directions within the window, and the minimum value of 4 as the angle corresponding function value of the target pixel. If the value is greater than the preset threshold, then the point is the corner of the image.

The specific approach of image corner detection algorithm is: assume pixel (x, y) as center of window w, and the window w moving Δx in the direction of x, moving Δy in the direction of y. Its local autocorrelation function is defined as follow:

$$
\begin{aligned}
m(x,y) &= \sum_{w}\left(-\left[I_x(\mathrm{x}_i,\mathrm{y}_i) \quad I_y(\mathrm{x}_i,\mathrm{y}_i) \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \\
&= \left[\Delta x \quad \Delta y \right] \\
&\begin{bmatrix} \sum_w I_x^2(\mathrm{x}_i,\mathrm{y}_i) & \sum_w I_x(\mathrm{x}_i,\mathrm{y}_i) I_y(\mathrm{x}_i,\mathrm{y}_i) \\ \sum_w I_x(\mathrm{x}_i,\mathrm{y}_i) I_y(\mathrm{x}_i,\mathrm{y}_i) & \sum_w I_y^2(\mathrm{x}_i,\mathrm{y}_i) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \\
&= \left[\Delta x \quad \Delta y \right] M(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}
\end{aligned} \tag{12}
$$

where, W is Gaussian window, which coefficient is $e^{-(x^2+y^2)(2\sigma^2)}$, $I_x(x_i, y_i)$ and $I_y(x_i, y_i)$ representation the gray change values in the direction of x and the direction of y. Gray matrix M contains a local neighborhood structure, which eigenvalues is a first order curvature of the autocorrelation function. If the both eigenvalues are high, then this point can be considered as a corner feature point.

$$
M = \begin{bmatrix} \sum_w I_x^2(\mathrm{x}_i,\mathrm{y}_i) & \sum_w I_x(\mathrm{x}_i,\mathrm{y}_i) I_y(\mathrm{x}_i,\mathrm{y}_i) \\ \sum_w I_x(\mathrm{x}_i,\mathrm{y}_i) I_y(\mathrm{x}_i,\mathrm{y}_i) & \sum_w I_y^2(\mathrm{x}_i,\mathrm{y}_i) \end{bmatrix} \tag{13}
$$

Suppose λ_1, λ_2 are the two eigenvalues of the matrix, corresponding to the three cases in shifting windows:

(1) In the flat region of the movable window, λ_1 and λ_2 are small, $m(x, y)$ changes small in the value;

(2) At the edges, one of λ_1 and λ_2 has a large value;

(3) At corner, the values of both λ_1 and λ_2 are large. Translations in any direction, the value of $m(x, y)$ will increase.

In practical applications, the corner can be determined by calculating the determinant and the trace matrix. The matrix does not need exact eigenvalues of M. The formula responses corner detection is as follow:

$$
H = \det M - k \cdot \left[tr(M) \right]^2 \tag{14}
$$

where, k is empirical value, which is usually taken from 0.04 to 0.06.When H is greater than the preset threshold T, the point is the corner.

$$
D = \iint_w \left[I_2\left(T(x,y) - I_2(x,y) \right) \right]^2 w(x,y)\,dxdy \tag{15}
$$

The standard method adopts $g_x = [-1,0,1;-1,0,1;-1,0,1]$ and $g_y = g_x^T$ as Harris gradient operators to calculate the image gradients of direction x and direction y. This calculation is simple, easy to implement, but it will be easily affected by noise. This paper improves gradient operators as follows: $g_x^{'} = [-1,2,1;-1,2,1;-1,2,1]$ and $g_y^{'} = g_x^{'T}$.

By increasing the weight 2, which indicates the importance of the center to achieve some degree of smoothing effect, it can achieve noise suppression.

The OpenCV algorithms calculate steps are as follows:

1) Use function CvMat( ) to define the corresponding matrix and image pointer，and then translating initial image to grayscale. For Each point of regional sub-images, according the horizontal differential operator $g_x^{'}$ and the vertical directions differential operator $g_y^{'}$ to calculate the gradient of direction x and y as well as the product, then, it can obtain five new images, which each pixel property values represent I_x，I_y，I_{xx}，I_{xx} $I_x I_y$.

2) Use a Gaussian template 5 * 5 to filter the five Gaussian images. Constructing matrix M, and according the formula (14) to calculate each pixel corner response.

3) Set the threshold $T = 8000$ and extract $H(x, y) > T$ corners as an alternative point.

4) Selecting the local extreme point. Feature points are localized within maximum response values of corresponding pixels. After performing the threshold value extraction, it is need to extract the image point of maximum value for all local interest. Extracting the maximum value successively from the center of each pixel in the 3 * 3 window, if the value of the center pixel interest point is the maximum, then the point is the feature point.

5) Applying OpenCV function to detect corners: cvMinMaxLoc ( ) function is used to extract the

maximum and minimum value from images or a set of data, as well as its location.

3.2 Normalized Cross Correlation

NCC (Normalized Cross Correlation) is a matching algorithm, which is based on the grayscale image and has been widely used in image stitching filed. Implementation of the algorithm is relatively simple. This algorithm is adapted to the linear variation of gray values, anti-white noise ability and stable of Image matching performance.

Set the window $T(w)$ in image I_1 as the corresponding area of the window w in image I_2. The similarity measure between two images based on NCC as follows:

$$S=\frac{\iint_w ABw(x,y)dxdy}{\sqrt{\iint_w Aw((x,y))dxdy}\sqrt{\iint Bw(x,y)dxdy}} \quad (16)$$

$$A=\left(I_1\left(T(x,y)\right)-\overline{I_1}\right) \quad (17)$$

$$B=\left(I_2\left(T(x,y)\right)-\overline{I_2}\right) \quad (18)$$

When the similarity measure is closer to 1, it means the images matching more accurate. NCC algorithm is adopted for rough matching corner to find all possible corners in the image I_2 corresponding to images I_1. For each corner $p(u_1,v_1)$ in the image I_1, selecting a size of $(2n+1)\times(2n+1)$ rectangle search area based on the point $p(u_1,v_1)$ as the center. Then calculate correlation coefficient for each of the given window between the corner p of the image I_1 and the corner q within the search window in the image I_2.

$$C(\mathrm{p,q})=\frac{\sum_{i=-n}^{n}\sum_{j=-m}^{m} M\times N}{(2n+1)(2m+1)\sqrt{\sigma^2(I_1)\times\sigma^2(I_2)}} \quad (19)$$

$$M=I_1(u_1+i,v_1+j)-\overline{I_1(u_1,v_1)} \quad (20)$$

$$N=I_2(u_2+i,v_2+j)-\overline{I_2(u_2,v_2)} \quad (21)$$

where, the average value of the point (u,v) of the image $I_k\,(k=1,2)$ is:

$$\overline{I_k(u,v)}=\sum_{i=-n}^{n}\sum_{j=-m}^{m}\frac{I_k(u+i,v+j)}{(2n+1)(2m+1)} \quad (22)$$

where, the variance of $(2n+1)\times(2n+1)$ areas at the point (u,v) of the image $I_k\,(k=1,2)$ is:

$$\sigma(I_k)=\sqrt{\frac{\sum_{i=-n}^{n}\sum_{j=-m}^{m} I_k^2(u,v)}{(2n+1)(2m+1)}\overline{I_k(u,v)}} \quad (23)$$

The range of correlation coefficient C is $[-1,1]$. The two relevant windows is dissimilarity when $C=-1$. The two relevant windows identical when $C=1$.

Given any corner in image I_1, then search the window area in the image I_2 to obtain the corresponding point which has the largest correlation coefficient. This point corner is considered as matching point for the given corner point in the image I_1, and then, a matching point set S_1 can be obtained. Similarly, given any corner in the image I_2, search the window area in the image I_1 to obtain the corresponding point which has the largest correlation coefficient. This point corner is considered as matching point for the given corner point in the image I_2, and then, a matching point set S_2 can be obtained. Finally, get the same match corners in two match points set S_1 and set S_2 .The corner is corresponding matched, completed the initial corner matching.

3.3 RANSAC

Draw on the experience of RANSAC algorithm and combining affine transformation model, a simplified method for screening the homonymy points is proposed for images registering[5]. After NCC, candidate of the homonymy points can be obtained, respectively $\{(x_1,y_1),(x_2,y_2),\dots,(x_p,y_p)\}$ and $\{(x_1',y_1'),(x_2',y_2'),\dots,(x_p',y_p')\}$. The corresponding matching similarity measure is $\{V_1,V_2,\dots,V_p\}$, then the steps of screening the homonymy points based on RANSAC as follows :

1) In the sequence of similarity measure, select three pairs of the homonymy points coordinates $\{(x_{a1},y_{a1}),(x_{a2},y_{a2}),(x_{a3},y_{a3})\}$ and $\{(x_{b1},y_{b1}),(x_{b2},y_{b2}),(x_{b3},y_{b3})\}$, which are correspond to the maximum 3 values of the relevant measure.

2) According to affine transformation model about images, using the above three points calculate parameters;

3) For the interest points set obtained by NCC from an image to be registered, according the affine transformation equations to obtain coordinate values

$\left\{\left(x_1^{"}, y_1^{"}\right), \left(x_2^{"}, y_2^{"}\right), \cdots, \left(x_p^{"}, y_p^{"}\right)\right\}$ of the homonymy points of the reference image;

4) For the set of points $\left\{(x_1, y_1), (x_2, y_2), \ldots, (x_p, y_p)\right\}$ and $\left\{\left(x_1^{"}, y_1^{"}\right), \left(x_2^{"}, y_2^{"}\right), \cdots, \left(x_p^{"}, y_p^{"}\right)\right\}$, according to the Euclidean distance formula(22), respectively calculated the distance of each homonymy points. If $d\left(\left(x_i, y_i\right), \left(x_i^{"}, y_i^{"}\right)\right) \le T_2$ the points consider as homonymy points, otherwise removed the points;

$$d\left(\left(x_i, y_i\right), \left(x_i^{"}, y_i^{"}\right)\right) = \sqrt{\left(x_i - x_i^{"}\right)^2 + \left(y_i - y_i^{"}\right)^2}, \quad 1 \le i \le p \tag{24}$$

5) Continue screening, till all homonymy points obtained from NCC has been processed.

3.4 Images stitching

Stitching several images is a process that fixing the coordinate of one image, and then translating the coordinate of the other images into the fixed image[7]. Let I_i be the point set of the i^{th} image, that is, I_1 is the point set of the first image. Fix the coordinate of the first image, and then translate the coordinate of the other image into the coordinate of the first image. For arbitrary I_i,

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \frac{1}{\lambda^{'}} * \begin{pmatrix} m_0^i & m_1^i & m_2^i \\ m_3^i & m_4^i & m_5^i \\ m_6^i & m_7^i & 1 \end{pmatrix} * \begin{bmatrix} x_i^{'} \\ y_i^{'} \\ 1 \end{bmatrix} \tag{25}$$

Unknown numbers $\lambda^{'}$、m_0^i、m_1^i、m_2^i、m_3^i、m_4^i、m_5^i、m_6^i、m_7^i are calculated through the matching points of the i^{th} image and the $i-1^{th}$ one. In this way, the point set of the i^{th} image is translated into the point set of the $i-1^{th}$ one. The result goes to that the point set of the i^{th} image is translated into the point set of the first image. Operate this to each image, translating point set of all images into point set of I_1 coordinate and realizing stitching of images.

3.5 Method of Image Fusion

Weighted average doesn't suit with the stitching of rotation Angle. This paper uses an improved weighted method. The basic method[6] is as follows:

1) Find the edge of the overlapping area, and tell the image edge. Define the two edges as edge 1 and edge 2.

2) Respectively calculate the minimum distance from point $P(x, y)$ of the overlapping area to edge 1 or edge 2. After calculation, the distance in image 1 is m and in image 2 is n.

3) The weighted coefficients are $\beta_1 = n/(m+n)$、$\beta_2 = m/(m+n)$. That is, the overlapping area can be expressed as follow:

$$I_0(\mathrm{x}, \mathrm{y}) = \beta_1 I_1(\mathrm{x}, \mathrm{y}) + \beta_2 I_2(\mathrm{x}, \mathrm{y}) \tag{26}$$

4 Experiment Results And Analysis

This experiment is operated in Windows 7, VS2010 software platform environment, combined with the use of OpenCV function library. Five image sequences, which size is 616*807 and with overlapping area, are used as experimental data for automatic stitching experiment. The overlapping area account about 30%. High-quality panoramic image will be generated because that the algorithm has good stability to light and contrast.

Figure 1 contains five corrected original pictures from five different cameras collecting at the same time.

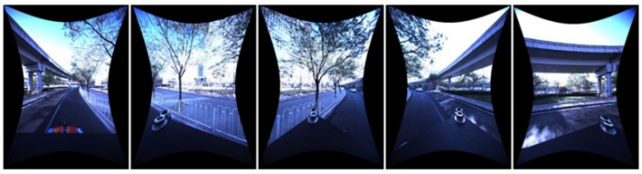

Fig. 1: （a）、（b）、（c）、（d）、（e）

Figure 2 displays five original images after cylindrical projection.

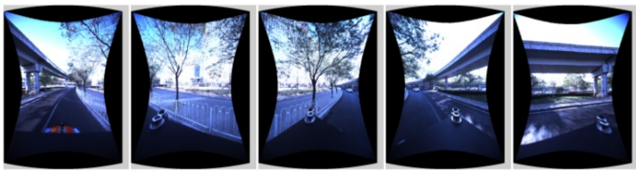

Fig. 2: （a）、（b）、（c）、（d）、（e）

Figure3 contains five cylindrical projection images after Harris corner detection.

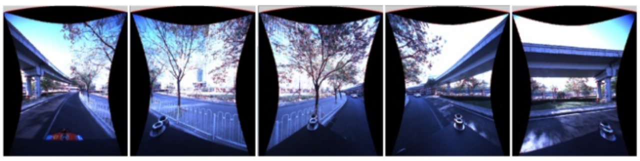

Fig. 3: （a）、（b）、（c）、（d）、（e）

Arbitral choose 2 of 5 images, which are overlapped. (this paper chooses image 1,2), to match images.

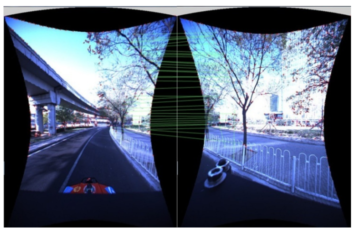

Fig. 4: corner matching

Figure 5 is Panoramic image Mosaic.

Fig. 5: panoramic image

From the perspective of the result of figure 5, for images of the multiple viewpoints and has a certain scale, the algorithm stitches images more naturally, without obvious seams. In order to further improve the quality of stitching, it is need to improve the accuracy of feature point positioning, feature point matching algorithm and image fusion which is based on wavelet in the future.

5 Conclusion

This paper studies an image auto-stitching method based on Harris corner feature. This algorithm has following characteristics: firstly, extract feature corner of images by using Harris corner detection and combine with NCC corner matching algorithm, which can reduce the amount of calculation, improve the location accuracy of corner, enhance the anti-noise performance of the algorithm and reduce the error response of bevel edges. Secondly, Projection inverse calculation is used when cylindrical projection transformed, to effectively correct distortion of images. Experiments show that this algorithm can not only reduce the amount of calculation, improve the speed, but also give consideration to the accuracy of registration stitching. It is greatly decreases the requirement of images overlapping area in stitching algorithm.

References

[1] Farzin. Mokhtarian, Riku. Suomela. Robust image comer detection through curvature scale space, [J] 1998(12).

[2] Cordelia. Schmid, Roger. Mohr Christian. Bauckhage. Evaluation of interest point detectors, [J] 2000(02).

[3] Chris. Harris, Mike. Stephens. A combined corner and edge detector, 1988

[4] Simone. Frintrop, Patric. Jensfelt, Henrik. I. Christensen. Pay Attention When Selecting Features, [J] 2006.

[5] Szeliski. R, Shun. H. Y. Creating. Full view panoramic image mosaics and environment maps, 1997.

[6] Hsieh. Jun-Wei. Fast stitching algorithm for moving detection and mosaic construction, 2004.

[7] Y. Xiong and K. Turkowski, Registration, calibration and blending in creating high quality panoramas, IEEE Workshop on Applications of Computer Vision, pp. 69-74, Oct. 1998.