Image Mosaic Based on Phase Correlation and Harris Operator

Fan YANG*, Linlin WEI, Zhiwei ZHANG, Hongmei TANG

School of Information & Engineering, Hebei University of Technology, Tianjin 300401, China

Abstract

In order to improve the insufficiency of harris corner, we present an auto-adjusted algorithm of image size based on phase-correlation. First, we detect the zoom relationship and translation co-efficiency between the images and modulate the unregistrated image’s scale to the same level as the original image. We obtain the Region of Interest (ROI) according to the translation parameter and then pre-treat the images and mark the interest points in the area by using improved harris corner operator. Secondly, we adopt Normalized Cross-Correlation (NCC) to wipe out the mismatched points preliminary after edging process, and get the final precise transformation matrix through Least Medium Square (LMedS) method. At last, we use method of weighted average to obtain a smooth mosaic image. The experimental results have shown that the setting of ROI and handling of the edge could cut the time down to about only half of the time consuming compared to SIFT; Besides, the scale difference between the images could enlarge from 1.8 to 4.7 and can eventually obtain a clear and stable mosaic result.

Keywords: Image Mosaic; Phase Correlation; Corner Detection; Image Registration

1 Introduction

Image mosaic is to composite two or more overlapped images into a seamless wide-angle image through a series of processing and is widely used in remote sensing [1], military matters [2] and medicine [3], etc. Due to the differences in rotation, exposure and location, etc. When taking these photos, it’s difficult to make a precise registration.

At present, methods based on the area, based on the frequency domain and based on the characteristics are the common registration methods we used. Area based method uses the most similarity principle among images to get the parameter by calculating the cost function (such as Sum of Squared intensity Differences [4], mutual information). This leads to a complex calculation involving lots of information and it’s very sensitive to the slight distinction between images. Method based on the frequency domain transforms the image from spatial domain to frequency domain, and get the relationships of translation, rotation and zoom through Fourier transformation (such as phase-correlation [5], Walsh transform [6]). But the use of this method

*Corresponding author.
Email address: yangfan@hebut.edu.cn (Fan YANG).
is very limited in particular application for requiring a big proportion of overlap area between the images to be mosaiced. A representative algorithm about feature extraction is registration based on Scale Invariant Feature Transform [7] (SIFT) brought by David G. Lowe. It’s perfect to dispose images with translation, rotation, zoom and illumination changes. But the description with 128-dimensions enlarges the calculation and cost of time to a great extent. PCA-SIFT [8] can shorten the mosaic time, but the decrease of dimensions would weaken the uniqueness of every vector, and we should calculate much more than SIFT. Harris corner detector [9] is obtained by improving Moravec operator. It’s robust with the changes in rotation, illumination, visual angle and noise, and it has a high speed because of the simple calculation method. However, it isn’t correct when disposing images with larger scale change so the application is very limited.

According to the comparison and analysis above, aiming at the mosaic between images that have larger scale difference, we try to synthesize the advantages both in frequency dispose and registration with features. A new robust method combined the phase-correlation and Harris corner is proposed. We can get the factor of translation and zoom by cross-power spectrum in order to optimize the detection of Harris. The feature detection then can be restricted in the overlapped area to avoid the waste of resource in irrelevant area when we do the search work. More importantly, this method can eliminate the non-adaptive weakness because of scale change. It is superior to SIFT and original Harris algorithm in terms of the calculation speed and applicability.

2 Image Scale Adjustment

2.1 Phase correlation algorithm

Phase correlation algorithm uses the cross-power spectrum to registration images and is used to get the translation factor initially. Imagine there are two images $I_1$ and $I_2$, and the translation between them is as following:

$$I_2(x, y) = I_1(x - x_0, y - y_0)$$

(1)

The Fourier transformation:

$$F_2(u, v) = F_1(u, v) \cdot e^{-j(ux_0 + vy_0)}$$

(2)

$F_1$ and $F_2$ are the Fourier transformation of $I_1$ and $I_2$. The cross-power spectrum is:

$$\frac{F_1 \ast (u, v) F_2(u, v)}{|F_1 \ast (u, v) F_2(u, v)|} = e^{-j(ux_0 + vy_0)}$$

(3)

$F_1 \ast (u, v)$ is the conjugate function of $F_1(u, v)$. We can get an impulse function $\delta(x - x_0, y - y_0)$ about the value of translation invariant $x_0$ and $y_0$ by using Fourier inverse transforming to formula (3). Fig. 1 is an experiment to get the range shift of lena by using phase correlation algorithm.

The preset value between image A and B is (20,50). There is an obvious impulse at the range shift coordination compared to the flat site on the inverse conversion curve. So we can get the shift value via the location of the maximum value.
2.2 Calculation of relative scale

For images that have relative relationships in location and scale, we can also get the zoom factor and rotation angle through a series of coordination transform \([10]\). For instance, two images \(f_1(x, y)\) and \(f_2(x, y)\) with a zoom factor \(\sigma\) and counterclockwise rotation angle \(\theta_0\):

\[
f_2(x, y) = f_1[\sigma^{-1}(x \cos \theta_0 + y \sin \theta_0), \sigma^{-1}(-x \sin \theta_0 + y \cos \theta_0)]
\] (4)

Looking for the polar former expression \(f_{1\rho}(\rho, \theta)\) and \(f_{2\rho}(\rho, \theta)\) about the images:

\[
f_{2\rho}(\rho, \theta) = f_{1\rho}(\rho/\sigma, \theta - \theta_0)
\] (5)

Hence we take the logarithm in the direction of \(\rho\) respectively with \(f_{1\rho}(\rho, \theta)\) and \(f_{2\rho}(\rho, \theta)\):

\[
f_{2\rho\lambda}(\lambda, \theta) = f_{1\rho\lambda}(\lambda - \lambda_0, \theta - \theta_0)
\] (6)

With the substitution \(\lambda = \log \rho, \lambda_0 = \log \sigma\). We can see in formula (6) that after the coordination transformation from cartesian to log-polar, the zoom changes and rotation will individually correspond to the shift in the direction of coordinating \(\rho\) and \(\theta\). Then we proceed the phase correlation part to get the factors \(\lambda_0\) and \(\theta_0\), while \(\sigma = e^{\lambda_0}\).

In order to emphasize the role of high-frequency component, we apply high-pass flow filter on the frequency spectrum before the FFT transform,

\[
H(m, n) = (1 - x(m, n))(2 - x(m, n))
\] (7)

with \(x(m, n) = \cos(\pi m) \cos(\pi n)\), and \(-0.5 \leq m, n \leq 0.5\). The algorithm is outlined in Fig. 2.

3 Feature Extraction

3.1 Extraction of Harris corner

In this paper, we get the corner response function by the ratio from the det and trace of matrix \(M\), which can avoid the randomness when choosing the scale factor compared to the method based on the difference value of the above ones. Besides, just as the experiments show, we could get much more stable features along with a speed-up procedure.

\[
R = \frac{\text{Det}(M)}{\text{trace}(M)} = \frac{\langle I_x^2 \rangle \cdot \langle I_y^2 \rangle - \langle I_x I_y \rangle^2}{\langle I_x^2 \rangle + \langle I_y^2 \rangle}^2
\] (8)
The non-maximum restrain of $R$ can be obtained with the help of a maximum filter in a window that covers the area with a radius of 3, and the point whose $R$ is less than the preset threshold would be removed.

3.2 Modified method of feature detection

Although brief and fast, the original Harris corner detection method has some shortcomings: even though it is robust to the illumination changes and rotations, it is very sensitive to the variation of image size. In addition, when doing a direct corner checking to images whose textures are dense or who have abundant details, we surely would get duplicate features in a local area. Inevitably, we must do extra work to extract and registration the points, including the useless ones. What’s more, we cannot ensure the quality of the estimated results. Besides, knowing that the corners are in accordance with the curvature of the edges, we have reasons to assume that all the valid candidate features are included in the edges we get. So additional preprocessing the image before extraction can offer a possibility to get more stable features.

Considering the analysis above and the work we have accomplished in chapter 2, we improve the extraction part as following:

Step 1: Get the shift and zoom factors with the help of phase correlation calculation.

Step 2: Modulate the unregistration image according to the zoom factor obtained from step 1 to get a couple of images with the same size. Then the images would have close texture distribution, which, in another word, expending the serviceable range greatly.

Step 3: Ascertain the ROI (Region Of Interest) between the images where they have the
overlapped area according to the shift factor we’ve got and the following work carries through just in this area. This leads to additional speed-ups and guarantees the efficiency of the transformation matrix when matching.

Step 4: Preprocess image before other works. The edge detection can reduce the search area and can greatly cut the matching-time down.

4 Estimation of the Homography

Not all the matches we get above are the correct ones. In this paper we use the Least Medium Square (LMedS) method [11] to obtain the inline points and transformation matrix. Two advantages are worth to mention comparing to the RANSAC (Random Sample Consensus) method [12] here: Firstly, LMedS doesn’t need to distinguish the inline and outline points, thus no need to preset the contrast threshold and this surely reduces the manual intervention; Secondly, number of interactions when using RANSAC is determined by the run time and we don’t know it exactly until after the procedure while LMedS doesn’t have this problem.

LMedS is also a kind of random parameter estimation method. It selects a subset of sample from the set of points randomly, and uses the Least Square (LS) to compute the model parameter, and then figures out the medium offset of all the samples and the model parameter obtained from this one.

Suppose the probability of any match that was a correct one in the data set is $\varepsilon$, thus by aiming at the $s$ pairs of matches we need to calculate the homography, the probability that at least once is the exactly match among $N$ times is

$$P = 1 - (1 - \varepsilon)^N$$

and then we can get $N$

$$N = \log(1 - P) / \log(1 - \varepsilon^S)$$

The detail procedures described as following:

Step 1: Confirm the times of sampling $N$.

Step 2: Fetch any random subset ($s = 4$) to get the transformation matrix $H$, and calculate the medium value about all features to the sum of error:

$$E_{med} = \text{med}_{i=1,2,...,n} \{d^2(x_{i,2}, Hx_{i,1}) + d^2(x_{i,1}, H^T x_{i,2})\}$$

$x_{i,1}$ and $x_{i,2}$ are the corresponding feature points while $d$ is the Euclidean distance.

Step 3: Repeat to execute step 2 for $N$ times and mark $H$ when the homologous $E_{med}$ is the minimum one.

The efficiency of LMedS is really poor when the noise of data is Gaussian distribution. In order to lower the bad influence, we need to calculate the standard deviation first:

$$\hat{\sigma} = 1.4628|1 + 5/(n - s)|\sqrt{E_{med}}$$

The matches whose $r_i^2 \leq (2.5\hat{\sigma})^2$ are all correct ones. By doing this we could remove most of the mismatched points and obtain the global optimum parameter estimation.
5 Experiments and Discussions

All the experiments we do were on the 32-bit computer with XP system, MATLAB 7.0 platform and the pictures involved were taken from lena photos, landscape and buildings. The goal of experimentation is to evaluate our approach and compare the mosaic results during the feature extraction and matching by different methods.

5.1 Results of the relatively scale by phase-correlation

The problem we focus on is to solve the scale non-invariant about Harris corner. Still with lean image as an example, the experiment was performed using the preset scales reported in Table 1, with a range variation from 1 to 4.5.

<table>
<thead>
<tr>
<th>Default scale</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
<th>4.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculate result</td>
<td>1</td>
<td>1.4845</td>
<td>1.9509</td>
<td>2.4472</td>
<td>2.8122</td>
<td>3.3128</td>
<td>3.7509</td>
<td>4.2383</td>
</tr>
</tbody>
</table>

Fig. 3 directly shows the measurement error of the two pairs of data as a form of line graph. Combined with Tab. 1 we can see that they are rather approximate to each other and the extent of the error is limited within 0.07 when the default scale is 1 to 4.7, even with a growth trend along with the increase of the preset value.

5.2 Comparison of matching results

The zoom factor between the images we show is 3.2. Number of candidate points and matching pairs is shown in Tab. 2, including both before and after the phase correlation process. As it shows, even though there are more corner points after direct detection, we get only 5 pairs of matches after the registration, which are unstable result either and it’s impossible to obtain a
satisfying and seamless result. However by using the method we proposed, we guaranteed the precise matching by selecting after the scale transformation that the number of reliable feature that matches is 56.

![Table 2: Comparison of the interest points detection](image)

<table>
<thead>
<tr>
<th>Number of point</th>
<th>Interest points</th>
<th>Matching points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original image</td>
<td>121</td>
<td>—</td>
</tr>
<tr>
<td>Direct detection</td>
<td>612</td>
<td>5</td>
</tr>
<tr>
<td>Detection after translation</td>
<td>96</td>
<td>56</td>
</tr>
</tbody>
</table>

Fig. 4 shows the comparison of images that we got during the procedure performed on lena photos that have 3.2 times of scaling: find the ROI first through phase compute, then mark the features just in this area as Fig. 4 (a) and Fig. 4 (b) show. Fig. 4 (c) is the mosaic result using the original Harris matching, which is seriously distorted for not able to conquer the obvious difference of scale. And we can obtain a distinct and accurate result by employing our method just as Fig. 4 (d) shows.

![Fig. 4: Comparison between real examples (scale=3.2)](image)

We have done the registration among about 30 pieces of photos aiming at characters, landscapes and structures, et al in our reality work. Fig. 5 and Fig. 6 are another two results about different scaling value. In Fig. 5 we can get the satisfactory mosaic image using both two methods, where the images with 1.6 times scale difference, however it’s obvious that the result of our method is much better than direct matching. The zoom factor is 4 in Fig. 6 and it’s obvious that the matching result is madly distorted, ended with surely a failure. Let’s then see the result Fig. 5 (d) shows, with no doubt that it’s successful. In conclusion after numerous tests we come to a conclusion that matching based on Harris corner only works when the zoom factor between the images below 1.8 and that’s the extremity and distortion has already emerged then. Through the calculation of phase correlation and adjustment to the scale relationship of the images, in contrast, we can solve the problems even the scaling is 4.7, and within this figure of course. Not only this leads to a much bigger scope of application for Harris, but also it is a promising mosaic result doing the matching work on the same order of scale.
5.3 Analysis of effectiveness

We have mentioned that the most valuable advantage of Harris operator compared with SIFT is its short running time. We hope this superiority could maintain even after the scale transformation, so we recorded the elapsed time of our method and confirmed the proposal afterwards.

Still with the above three groups of images as an example, from Tab. 3, we see the contrast of time-consuming between SIFT arithmetic and method used in this paper. Direct at images with diverse scaling, as expected, we only used nearly half of the time that SIFT costs to get the equal satisfying results. The explanation is as following: Although we added the amount of work for phase section, we delimit the ROI. And along with the preprocessing it decreased the workload dramatically. As a result the time-consuming was decreasing instead of increasing. This superiority can be performed better when we conduct the mosaic work aiming at big scenes and panoramic images.

<table>
<thead>
<tr>
<th>Mosaic method</th>
<th>SIFT method time/s</th>
<th>Our method time/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>lena</td>
<td>4.859</td>
<td>2.641</td>
</tr>
<tr>
<td>Landscape</td>
<td>4.173</td>
<td>1.942</td>
</tr>
<tr>
<td>Structure</td>
<td>5.683</td>
<td>3.927</td>
</tr>
</tbody>
</table>
6 Conclusion

In this paper we have presented an approach for image mosaic based on phase-correlation and Harris operator. First the scaling and translation relationship is gained according to the correlation known by phase-correlation. Then the unregistered image is adjusted and the ROI scope of matching is limited all according to the factors derived above. Finally the feature points are detected and matched just in this area, based on the improved Harris corner. We comprehensively apply the advantages of spatial and frequency domain to conquer Harris’s maximum inadequacies for not possessing the scale-invariant quality, and robustness is enhanced too. What’s more, the setting of ROI and adoption of preprocessing avoid the useless extraction and registration thus leading to additional speed-ups and improvement of the precision. We have exhibited simulations and the actual results tell that the scaling permitted between the images enlarged from 1.8 to 4.7 while the relatively time-consuming is about one half compared to SIFT, under the condition that the mosaic result is high-quality.

References