A Coarse-to-Fine Registration Method Based on Geometric Constraints of Block and Parallel Architecture

Zetao Jiang1, 2, Chuan Guo2
1Guilin University of Electronic Technology, Guilin, China
2Nanchang Hangkong University, Nanchang, China
1zetaojiang@126.com, 1nchuguochuan@126.com

ABSTRACT: Aiming at long registration time and mismatch problems resulted by images that exist local region similar or generated SIFT vector similar in the traditional SIFT registration method, a fine registration method based on geometric constraints of block and parallel architecture is put forward. This method using SIFT algorithm parallelly extract feature points to calculate initial transform matrix, in order to provide geometric constraints for block and fine registration, after segment the overlapping region into several blocks, we do blockwise SIFT matching using parallel architecture to achieve fine registration. During the fine registration, we use affine invariance of Mahalanobis distance to screen feature points and eliminate duplicate matches and mismatches. The experimental results show that this method eliminates the mismatch generated by same local features, registration accuracy and speed has improved, the proposed approach has practical value.

Keywords: From Coarse To Fine, Geometric Constraints, Affine Invariant, Block Strategy, Parallel

Received: 10 October 2015, Revised 9 November 2015, Accepted 15 November 2015

© 2016 DLINE. All Rights Reserved

1. Introduction

Image registration and fusion splicing technology is the hot issue of the image processing, machine vision and computer graphics research. Registration have a wide range of applications in image fusion, video monitoring, automatic target recognition, biomedical analysis, field of virtual reality, military, and remote sensing technology, such as other fields. Image matching is putting some images of the same scene with overlap sequence images into a high resolution image with big size and wide angle of view. Image registration [1,2,3] is the key technology in the process of image mosaic, it affects the precision and speed of image stitching. Currently the most widely used registration method is based on feature points matching [4-9, 21], a method that extract the feature points in the images through implementing feature detection to paired image to be registered, then complete image registration after generating feature descriptors.

C.Harris and M.J.Stephens [10] proposed the famous Harris operator. Harris corner is invariant to noise, rotation and illumination, but it is sensitive to scale changes, so it was not used in multi-scale image registration. Lowe [11] proposed a landmark local feature based on the theory of scale space, i.e., the famous SIFT (Scale Invariant Feature Transform) operator. SIFT is invariant to scale, rotation, transformation, and illumination changes of image. But the dimensions of the SIFT feature descriptor is higher, it is not applicable to real-time demand higher occasion. Later, SURF [12,13], SSIFT [14] and PCA-SIFT[15,16] algorithms are
effective for accelerating SIFT, these algorithms improve the efficiency of the registration, but the accuracy is worse than SIFT algorithm. Accordingly we choose SIFT points to get initial matching. The experimental results show that this method produces some mismatch problems which resulted by local region similar or SIFT vector similar although the SIFT feature point accuracy is higher, as shown in Figure 1 use red box marked. Ji Hua[17] who proposed a combination of global information SIFT feature matching algorithm, to a certain extent, solved the mismatch problem caused by local region similar. However, this method increases the amount of calculation and it is sensitive to scale changes. Bastanlar [18] proposed an improved SIFT algorithm to solve the large scale difference problem of image registration and improved registration accuracy. However, the method does not solve the above problem of mismatch. Goncalves [19] proposed a image registration method based on the combination of image segmentation and SIFT to eliminate the mismatch, however, this method cannot be applied to the time response to demand higher occasion.

![Figure 1. Mismatch generated by same local features](image)

This paper proposed a coarse-to-fine registration method based on geometric constraints of block and parallel architecture, for mismatch problems which resulted by local region similar or SIFT vector similar, and spent too long time on registration. First, we use SIFT algorithm parallely extract feature points to calculate initial transform matrix in order to provide geometric constraints for block and fine registration, after segment the overlapping region of two images into several blocks, we do blockwise SIFT matching using parallel architecture to achieve fine registration. During the fine registration, we use affine invariance of Mahalanobis distance to screen feature points further and eliminate duplicate matches and mismatches to improve matching accuracy. Then, outlier removal[20] after finishing fine registration, transformation estimation, image resampling, and image mosaicing.

2. Multi-scale Feature Detection and Initial Matching

2.1 SIFT Point Detection

The SIFT algorithm first constructs Gaussian pyramid, the scale space $L(x,y,\delta)$ of image $I(x,y)$ will be constructed by convolution of image $I(x,y)$ and Gaussian kernel function. That is, $L(x,y,\delta)=G(x,y,\delta) \ast I(x,y)$, where, $G(x,y,\delta)=\frac{1}{2\pi\delta^2}e^{-\frac{x^2+y^2}{2\delta^2}}$, $\delta$ is the scale factor of Gaussian function, in order to control the level of image’s smoothness, we can change the size of the $\delta$. We intend to construct $O$ octaves $S$, levels images and appoint scale value $\delta$ for each image defined as below: $\delta(0,s)=2^{-1}\times k^s \times \delta_0$, where $\delta_0$ means initial scale value, $o$ and $s$ represent $0th$, $sth$ image. $k=2^{1/8}$ is Proportional constant of the scale factor.

Lowe[11] experimentally found detecting extreme points on the Difference of Gaussian scale-space can efficiently extract stable feature point candidates. Then, Gaussian DoG images will be obtained by DoG operator: $D(x,y,\delta)=[G(x,y,k\delta)-G(x,y,\delta)] \ast I(x,y)$. 
The process of SIFT algorithm to extract the feature points: first we construct Gaussian pyramid (O octaves, S levels images) and then we construct Difference of Gaussian space. Feature points are made up of 3-D local extreme points of Difference of Gaussian images. After detecting extreme points we can’t regard them as SIFT points until the heavy edge corresponding is eliminated through the Hessian matrix.

2.2 SIFT Feature Descriptor
Feature descriptor is used to describe the gradient statistics of the feature point neighborhood, it is unique. SIFT descriptor can be obtained by gradient histogram of 16*16 around the feature point. As show in Figure 2. In order to ensure the scale invariance of feature descriptor, we compute the feature descriptor on Gaussian blur image $L(x,y,\delta)$ that scale is equal to feature point. In order to guarantee the rotation invariance of feature descriptor, we compute the main orientation of feature point. And then we divide 16*16 into 4*4 small regions with the size 4*4. Next, we make a statistics about 8 gradient histogram of every 4*4 region and obtain the descriptor vector whose dimension is 4*4*8=128. In order to eliminate the influence of illumination differences, we normalize the vector. So far, the descriptor describes the change of gradient orientation around the feature point. Therefore it is robust to illumination, scale and rotation. Due to the elimination of edge corresponding, it is also robust to noise.

![Figure 2. The 16*16 window that computes feature point descriptor](image-url)

2.3 Initial Matching of Feature Points
After the detection and description of features, we utilize Euclidean distance and BBF(Best Bin First) to get initial matching pairs. Let $P_a$ and $P_b$ be SIFT point set to be matched, SIFT point set $P_a$ is derived from the image $I_L$ and SIFT point set $P_b$ is derived from the image $I_R$. We get initial matching by following steps:

1) Segmentation feature point set $P_b$ and construct complete K-D tree.

2) For one point $a_i (i=1,2...,n) (a_i \in P_a)$ of $P_a$, we find two feature points (Nearest Neighbor/Next nearest neighbor) $b_j (j=1,2...,m;j \neq k) (b_j \in P_b, b_k \in P_b)$ in K-D tree, and then we get the two distance $D_{nearest}$ and $D_{next\_nearest}$. 

3) Compute the ratio: $ratio = \frac{D_{nearest}}{D_{next\_nearest}}$, we set a ratio threshold $T_{ratio}$. If $ratio < T_{ratio}$, $a_i$ and $b_j$ are matched, add the pair of match points $[a_i, b_j]$ to the corresponding matching point set $Z$, go to step 4. Otherwise, Skip to step 2 after $i + 1$, compute the matching point of next feature point $a_{i+1} (a_{i+1} \in P_a)$.

4) Output the set of feature matching until all points of $P_a$ are computed. Otherwise, Skip to step 2 after $i + 1$.

3. SIFT Block Parallel Fine Registration Method
3.1 Regional Division and Geometric Constraints
The purpose of coarse matching is to calculate the initial matching transformation matrix $T$, it can provide geometric constraints for regional division and fine registration. This step does not need too many feature points, in order to improve the efficiency of computation. We downsampling the source images before doing the coarse registration. In order to preserve the image details as many as possible, the downsampling rate should not be too high. The downsampling rate is defined as follow:

$$
r = \min(2^n, 8), \quad n = \left\lfloor \log_2(M_s/M_d) \right\rfloor
$$

$$
M_s = \min(W_s, H_s), \quad M_d = \min(W_d, H_d)
$$

Where the $W_s$ and $H_s$ the are the width and height of the source image, the $W_d$ and the $H_d$ are the width and height of the image after downsampling.

According to based on the method of section II, we can implement coarse registration and calculate the initial transformation matrix $T$, then determine the overlapping area in the source image through basis transformation matrix and geometric constraint relation of matching points, then divide the overlapping area into blocks, at last, implement fine registration in the regional blocks parallelly. As shown in figure 3 (the left image is a reference image, the right image is a image to be registered). Firstly, determine the overlapping region in the reference image: we map the top left corner and the bottom left corner points ($c$ and $d$) of the right image to the left image, i.e. $c' = T(c), d' = T(d)$. By this way, we obtain the overlapping region ($R_{ref}$) of reference image. In the same way, we map the top right corner and the bottom right corner points ($a$ and $b$) of the left image to the right image through the inverse transformation of $T$, i.e., $a' = (\text{inverse}(T))(a), b' = (\text{inverse}(T))(b)$ and we obtain the overlapping region($R_{old}$) of source image(image to be registered). Here, these points ($c', d', a', b'$) could not on the edge of image, in order to facilitate subsequent processing, we adjust them to the edge coordinates.

![Figure 3. Determine overlapping regions of the reference image and the image to be registered](image)

After determine the overlapping area, we segment the region according to the size of overlapping area: (1) If the overlapping area is more than 80% of the original image, region division of $3 \times 3$ template is implemented on the original image pairs. (2) If the overlapping area is more than 40% and less than 80% of the original image, region division of $3 \times 2$ template is implemented on the overlapping area of original image pairs. (3) If the overlapping area is less than 40% of the original image, region division of $3 \times 1$ template is implemented on the overlapping area of original image pairs. This approach has two advantages: if registration image pairs has obvious non-overlapping region, this method not only saves memory space and reduces the unnecessary calculation and improves the efficiency. Secondly, we use constraint of block and geometric constraints of the initial transformation matrix to eliminate the mismatch generated by same local features to improve matching accuracy. As shown in figure 4, Assume that descriptors of these feature points($p_1, p_2, p_3, p_4$) are similar. For $p_1, p_2, p_4$ three feature points, Due to the influence of angle and noise, the Euclidean distance between $p_1$ and $p_4$ may be less than the Euclidean distance between $p_1$
and $p_2$, hence, traditional SIFT registration method causing a mismatch ($p_1, p_4$). The corresponding matching point of the feature point $p_1$ in Region $R_1$ may only appear in the corresponding region $R_2$, proposed method no longer search matching point in the region of $R_3$ through the constraint of region block, and the mismatch ($p_1, p_4$) is removed effectively. For $p_1, p_2, p_3$ three feature points, due to the influence of angle and noise, the Euclidean distance between and may be less than the Euclidean distance between $p_1$ and $p_2$, hence, traditional SIFT registration method causing a mismatch ($p_1, p_3$). We can eliminate the mismatch through geometric constraint which provided by the initial transformation matrix: firstly, we map the $p_3$ to the reference image through the initial transformation matrix $T$, then, calculate the distance between them, i.e. $\text{dis}(T(p_3), p_1)$, if $\text{dis}(T(p_3), p_1) \geq \tau$ (is a threshold), the mismatch ($p_1, p_3$) are removed successfully.

![Figure 4. Eliminate the mismatch through regional division and geometric constraints](image)

### 3.2 Parallel Architecture
Parallel computing can simultaneously handle multiple tasks or data, it can significantly improve the efficiency of the processing tasks. Multithreading is a simple and effective way to implement parallel computing. A process contains a number of different execution units (thread), one of them is the master thread, it is responsible for task allocation and management of the sub threads. When the program is running on multi-core systems, the main thread assigns tasks to each child thread after initialization data, then, sub threads are assigned to run on a different CPU, the master thread collects and processes information after all tasks are finished.

To further improve the efficiency of this method, we implement the downsampling and blockwise SIFT registration procedure in a parallel fashion. Firstly, in the coarse registration phase the following steps are implemented based on multi-threading: downsampling of the original image pairs (image $I_L$ and $I_R$), feature point extraction, feature descriptor generation. Then, the main thread collects feature point set and calculates the transformation matrix, after finishing coarse registration and segment the overlapping region into several blocks at the master thread, image block pairs are sent to the sub thread for blockwise SIFT matching. Finally, the master thread collects the matched features and computes the fine transform matrix after all tasks are finished.

### 3.3 Affine Invariant
While calculating the principal direction of SIFT descriptor, if the height of the gradient direction is bigger than or equal to 80% of the principal direction, the direction was added as a auxiliary direction. This feature points are copied and given the new direction, so, one feature point may have more than one direction, and there will be a several copies. In the process of matching, part or all of these feature points that above were mentioned will be become matching points. Actually, they are the same feature point, hence, such repeat points should be removed. Bi-direction matching can remove such repeat matching point. On the other hand, SIFT feature vector is a description of the neighborhood of feature point, it doesn’t take into account the distribution information of feature points and geometric information. Therefore, the matching that simply relies on the Euclidean distance
between two SIFT vector will appear a lot of mistakes. In figure 5, suppose that we finished the feature point that matching in source image pairs \( I_L \) and \( I_R \), three pairs of matching points were listed on here \([a, b], [a', b'], [a'', b'']\). As shown in figure 5, the two triangles are formed from three feature points of image \( I_L \) and \( I_R \) image respectively, Translation, rotation and scaling transformation exists between them, and the two triangles are affine equivalent, traditional SIFT matching Based on Euclidean distance do not consider this geometric constraints.

In order to eliminate repeat matching and mismatching that were caused by above-mentioned two reasons, in this paper, we uses the bidirectional matching. The first matching uses Euclidean distance, then, we can further screening the feature points through applying affine invariance of Mahalanobis distance in the second matching. In this paper, the bidirectional matching algorithm is as follows (the feature points set of \( I_L \) and \( I_R \) is \( P_a \) and \( P_b \)):

1). Search feature points through the BBF algorithm that mentioned in section 2.3, and similarity measure uses Euclidean distance. Find matching points in \( P_b \) corresponding to each feature points in \( P_a \) in proper order, so that we can obtain the initial matching points \( \text{Set}_{\text{first}} = \{ [a, b], [a', b'], [a'', b''] \} \). The Euclidean distance calculation formula is as follows:

\[
\text{Dis}_{\text{Euclidean}} = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}
\]

2). Filtering out the false matches through affine invariants of Mahalanobis distance. Findind the matching points in \( P_a \) corresponding to each feature points in \( P_b \) in proper order, then, the second matching points set is obtained \( \text{Set}_{\text{second}} = \{ [b, a], [b', a'], [b'', a''] \} \). Mahalanobis distance calculation formula is as follows:

\[
\text{Dis}_{\text{mahalanobis}} = \sqrt{(a - \mu)^T C^{-1} (a - \mu)}
\]

where \( a = (a_x, a_y) \) is an sample point in \( P \) and \( Q \) which are sample space feature point sets \( a \in P_a \), \( \mu \) is the average, \( C \) represents the covariance matrix, \( \mu \) and \( C \) are defined as follows:

\[
\mu = (\mu_x, \mu_y) = \frac{1}{n} \left( \sum_{i=1}^{n} a_{ix}, \sum_{i=1}^{n} a_{iy} \right)
\]

\[
C = \frac{1}{n} \left[ \begin{array}{cc} \sum_{i=1}^{n} (a_{ix} - \mu_x)^2 & \sum_{i=1}^{n} (a_{ix} - \mu_x)(a_{iy} - \mu_y) \\ \sum_{i=1}^{n} (a_{iy} - \mu_y)(a_{ix} - \mu_x) & \sum_{i=1}^{n} (a_{iy} - \mu_y)^2 \end{array} \right]
\]

3). Calculating the final set of matching points, \( \text{Set}_{\text{opt}} = \text{Set}_{\text{first}} \cap \text{Set}_{\text{second}} \).

After all, the concrete steps of this paper are:
1). Reduce the source image pairs to low resolution level through direct downsampling, the downsampling rate \( r = 2 \), and we
obtain the image pairs $I_1$ and $I_2$.

2). Feature Detection. The feature point sets $P_a$ and $P_b$ of image $I_1$ and $I_2$ are obtained from SIFT algorithm, and the process is based on multithreading.

3). Initial Matching and Calculate the initial transformation matrix $T$.

4). Apply Section 3.1 to determine the overlapping area of the source image pairs, and segment the overlapping region of two images into several blocks.

5). Extract the SIFT feature points in each block pair based on parallel architecture (multithreaded) described in section 3.2, apply the third section method to implement fine registration.

6). Collect the matched keypoints from each block pair and compute the final Transformation Matrix $T'$, then, obtain precise stitching images after resampling.

4. Experiments and Discussion

To assess the precision and the efficiency of the proposed approach in detail, two sets of experiments are designed. The first set of experiments is to verify the registration accuracy of this method, and the second set of experiments aims to show the registration speed of the proposed approach. The experiments are implemented in Matlab and run on Windows 7 system, the computer’s processor is Intel (R) Core (TM) 2 Quad CPU 2.83GHz, 4GB of memory itself. In this paper, the experimental use of original images shown in Figure 6:

![Figure 6. The source images](image_url)
In order to verify the block and geometric constraint information can effectively eliminate the mismatching caused by local region similarity, we compared the proposed approach with the traditional SIFT registration method (SIFT + Euclidean distance). In this experiment, we use the matching ratio to evaluate the effectiveness of the block and the geometrical constraint, the computation formula of matching rate is as follows: 

\[
R_{match} = \frac{N_{cm}}{N_m},
\]

where, \(N_m\) is number of matched, \(N_{cm}\) is number of correct matches. The experimental results are shown in figure 7, the matching rate of different approaches are listed in Table I.

Figure 7(a). matching result based on the traditional SIFT matching approach

Figure 7(b). Matching result based on the proposed approach.

From figure 7(a), we can see that many mismatches obviously. It is because of simply rely on the Euclidean distance between two SIFT vectors, images that exist local region similar or local region not similar but generated SIFT vector similar will result in a close Euclidean distance, thus, these feature points are mistaken for matching point pairs. The proposed approach did not appear this situation after coarse registration and segment the overlapping region. Additionally, we use affine invariants of Mahalanobis distance further screening the matching points and eliminate mismatches. Thereby, the registration precision is improved significantly. Comparison of the rate of matching are listed in Table I. The traditional SIFT matching approach can achieve 739 pairs, and 335 pairs are correct, while the proposed approach can achieve 568 pairs and 326 pairs are correct. The matching ratio increases from 45.3% to 57.4%. The comparisons indicate the advantage of the block and the geometrical constraint in eliminating the false matches.

<table>
<thead>
<tr>
<th>Traditional SIFT Matching Approach</th>
<th>Initial Matches</th>
<th>Correct Matches</th>
<th>Matching Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>739</td>
<td>335</td>
<td>45.3%</td>
</tr>
<tr>
<td>The Proposed Approach</td>
<td>568</td>
<td>326</td>
<td>57.4%</td>
</tr>
</tbody>
</table>

Table 1. Performance Comparison On Matching Ratio
To verify the registration efficiency of the proposed approach, we compare the execution time of following three methods: the proposed approach, the traditional SIFT matching approach based on the single thread, and the proposed approach that based on the single thread. The results shown in Table II.

<table>
<thead>
<tr>
<th></th>
<th>The Traditional Sift Registration Approach Based On The Single Thread</th>
<th>The Proposed Approach Based On The Single Thread</th>
<th>The Proposed Approach (Based On Multithreading)</th>
</tr>
</thead>
<tbody>
<tr>
<td>time(s)</td>
<td>2.1722</td>
<td>3.7212</td>
<td>1.3589</td>
</tr>
</tbody>
</table>

Table 2. Performance Comparison At Registration Time

From Table II, we can see that the registration efficiency of the proposed approach is improved significantly, and the execution time is reduced from 2.1722 to 1.3589. This comparison demonstrates the advantages of multithreading in improving the efficiency. The registration time of the proposed approach based on single thread is longer than the traditional SIFT registration method based on single thread. This is because the method spent more time in the following aspects: Coarse registration, regional division and using the geometric constraints and affine invariants screening the matching points. The proposed method makes full use of the advantages of parallel, and improves the efficiency of the registration significantly. The following processes are completed based on multi-threading: downsampling of the original image pairs, Establishment of a Gaussian pyramid, Feature point extraction, and the fine registration of each block pair.

We calculate the final transformation matrix through the matching point set after finishing fine registration, then, image resampling, the result of image stitching as shown in Figure 8.

Figure 8. Matching result Mosaic result by the proposed approach

5. Conclusion

This paper proposed a coarse-to-fine registration method based on geometric constraints of block and parallel architecture, for mismatch problems which resulted by images that exist local region similar or local region not similar but generated SIFT vector similar, and spend too long time on registration. First, we use SIFT algorithm parallelly extract feature points to calculate initial transform matrix in order to provide geometric constraints for block and fine registration, after segment the overlapping region of two images into several blocks, we do blockwise SIFT matching using parallel architecture to achieve fine registration. During the fine registration, we use affine invariance of Mahalanobis distance to screen feature points further and eliminate duplicate matches and mismatches to improve matching accuracy. The experimental results show that our method eliminates the mismatch generated by same local features, registration accuracy and speed has improved, the registration rate increased from 45.3% to 57.4% and time reduced from 2.1722s to 1.3589s. Our method also requires improvements in the following respects: template 2 * 3 used in overlapping regional division did not thorough explore how to divide the overlapping region to further improve the registration rate, and therefore we need to study how select the number of blocks self-adaptively according to the size of the
image resolution and the size of overlapping region. On the other hand, the improvement of the speed depends on the multithread, so we need to optimize algorithms to improve the efficiency of the algorithm itself.

Acknowledgements

This paper is sponsored by Nature Science Foundation of China (61272216) and innovation Foundation of aerospace in China (CASC201102).

References


46 (5), 346-348.

