J. Shanghai Jiao Tong Univ. (Sci.), 2017, 22(1): 99-106
DOI: 10.1007/s12204-017-1807-7

Algorithm Based on Morphological Component Analysis and Scale-Invariant Feature Transform for Image Registration

*WANG Gang** (王　刚), *LI Jingna* (李京娜), *SU Qingtang* (苏庆堂)
ZHANG Xiaofeng (张小峰), *LÜ Gaohuan* (吕高焕), *WANG Honggang* (王洪刚)
(School of Information and Electrical Engineering, Ludong University, Yantai 264025, Shandong, China)

Abstract: In this paper, we proposed a registration method by combining the morphological component analysis (MCA) and scale-invariant feature transform (SIFT) algorithm. This method uses the perception dictionaries, and combines the Basis-Pursuit algorithm and the Total-Variation regularization scheme to extract the cartoon part containing basic geometrical information from the original image, and is stable and unsusceptible to noise interference. Then a smaller number of the distinctive key points will be obtained by using the SIFT algorithm based on the cartoon part of the original image. Matching the key points by the constrained Euclidean distance, we will obtain a more correct and robust matching result. The experimental results show that the geometrical transform parameters inferred by the matched key points based on MCA+SIFT registration method are more exact than the ones based on the direct SIFT algorithm.
Key words: image registration, morphological component analysis (MCA), scale-invariant feature transform (SIFT), key point matching
CLC number: TN 911　　**Document code:** A

0 Introduction

Image registration is the process of geometrically aligning two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. Image registration is a crucial step in many image analysis tasks, such as in remote sensing, medical imaging and computer vision. The majority of the registration methods revolve around the two aspects: feature detection and feature matching. The features are the salient and distinctive objects which are frequently spread over the images and which are easily detectable in the two images, such as closed-boundary regions, edges, contours, line intersections, corners, etc. Simultaneously, the features are stable in time to stay at fixed positions during the whole experiment.

The detection methods should have good localization accuracy and should not be sensitive to the assumed image degradation. In an ideal case, the algorithm should be able to detect the same features in all projections of the scene regardless of the particular image deformation. Because the feature-based ones do not work directly with image intensity values, the features represent information on higher level. In most papers, the detection of line and point features is prevalent[1]. The line features can be the representations of general line segments, object contours such as coastal lines, roads, etc. Line correspondence is usually expressed by pairs of line ends or middle points. The common line feature detection methods include the Canny detector and the Laplace-Gaussian detector[2]. However, the thresholds of the detector need to be set artificially to ensure that the extracted object information is accurate in the complicated background. The point features can be the most distinctive points with respect to a specified measure method, for example, the line intersections, high intensity variance points, local curvature, and local extrema of wavelet transform etc[3]. The most well-known point feature is the Harris corner feature[4], which is understood as points of high curvature on the region boundaries. Much effort has been spent in developing precise, robust and fast methods for corner detection. Kitchen and Rosenfeld[5] proposed to exploit the second-order partial derivatives of the image function for corner detection. Dreschler and Nagel[6] searched for the local extrema of the Gaussian

Received date: 2016-03-12
Foundation item: the National Science Foundation of China (No. 61471185), the Natural Science Foundation of Shandong Province (No. ZR2016FM21), Shandong Province Science and Technology Plan Project (No. 2015GSF116001) and Yantai City Key Research and Development Plan Project (Nos. 2014ZH157 and 2016ZH057)
∗**E-mail:** happy_wg@163.com

curvature. Corners are widely used as point matching mainly because of their invariance to imaging geometry and because they are well perceived by a human observer. However, corner detectors based on the second-order derivatives of the image function are sensitive to noise. It is difficult to set the threshold and the size of the neighborhood to extract the corner points with the same property when illumination changes are expected. Moreover, the Harris corner detection algorithm is not invariant to the scale and rotation changes, it needs to design additional methods to match the corresponding corner points.

Considering the transformation (i.e. horizontal and vertical translation, scaling factor, and rotation angle) has taken place, we should find an appropriate method to achieve the point matching. Such a technique is based on scale invariant features transform (SIFT)[7] algorithm which is used to robustly detect and describe clusters of points belonging to cloned areas. Successively, these points are exploited to reconstruct the parameters of the occurred geometric transformation.

To summarize, the use of feature-based methods is recommended if the images contain enough distinctive and easily detectable objects. For example, the typical images contain a lot of objects which have sharp edge and angular features. We can find the correct correspondence of the key points in the different two images by the SIFT method. However, some images such as landscape images cannot easily and stably construct the key points because the important geometric features in these images are usually disturbed by the texture and noise, they will produce several mismatching points. Thus these images are inappropriate to apply the SIFT detection method directly. How can we solve this situation above? In the view of human perception process, the important information which will be easily detected is the profile of the objects in the image, e.g., the basic geometrical information. Inspired by this point, we think that the important work is to find the feature which can describe the local or the whole basic geometrical characteristics in the image. As we know, an image can be decomposed by some models[8] into its building parts, such as texture, cartoon parts and noise. In general, the cartoon part formed by homogeneous regions and with sharp boundaries and texture part is small scale repeated details. Then the cartoon part is exactly the geometrical information. And it is helpful to extract the key points in the next step. From another point of view, the cartoon part corresponds to the low-frequency information in the image, which is stable and unsusceptible to noise interference.

In 2003, Vese and Osher proposed a model for image decomposing by energy minimization and partial differential equations[9]. This model can decompose a given image f into the sum $u+v$, where $u \in BV$ is a bounded variation (a cartoon representation of f), and v is an oscillatory function, representing random noise or texture. However, the model cannot distinguish between noise and texture. In order to avoid the disadvantages, the morphological component analysis (MCA) algorithm has been proposed for image decomposition by Starck and Elad[10]. It is a fusion of the Basis-Pursuit (BP) algorithm and the Total-Variation regularization scheme. However, the algorithm has not thought how to devise the effective dictionaries for sparse representation. This will lead to higher memory requirement and computational complexity. In this paper, we use the perception mechanism of the human visual system to devise the appropriate dictionaries. The new formulation leads to improve the computation efficiency and direct the noisy image towards a successful separation. Thus we can lay a good foundation for the feature matching procedure.

1 Morphological Component Analysis

According to the Ref. [10], the morphological analysis model is shown as follows. For an input image f consisting of a sum of K building parts, a typical assumption is made that the given image is a linear mixture: $f=\sum_{i=1}^{K} f_i$. Each building part f_i represents a different morphological component. For every part f_i, there exists a dictionary φ_i, f_i can be sparsely represented in φ_i and not as sparse as in other φ_j $(j \neq i)$. It is written as

$$\left.\begin{array}{l}\min \|c\|_0 \\ \text{s.t. } f_i = c\phi_i\end{array}\right\}, \tag{1}$$

where $\|c\|_0$ denotes the L_0 pseudo-norm of the vector c (i.e. the number of non-zero coefficients of c). For the sake of the desired solution, the problem formulated in Eq. (1) is non-convex and hard to be solved. Donoho proposed the BP method[11] which suggests the replacement of the L_0 norm with L_1 norm, thus Eq. (1) can be rewritten as

$$\left.\begin{array}{l}\min \|c\|_1 \\ \text{s.t. } f_i = c\phi_i\end{array}\right\}. \tag{2}$$

Here we assume that the cartoon dictionary $\{\phi_n\}$ and the texture dictionary $\{\phi_t\}$ have fast transform operator $T(c_k = T_k f_k)$ and inverse transform operator $T^{-1}(f_k = T_k^{-1} c_k)$. So the morphological analysis model can be rewritten as

$$\begin{aligned}\{f_t^{\text{opt}}, f_c^{\text{opt}}\} = &\arg \min_{\{f_t, f_c\}} \|T_t f_t\|_1 + \|T_c f_c\|_1 + \\ &\mu \|f - f_t - f_c\|_2^2,\end{aligned} \tag{3}$$

where $f_t = T_t^{-1} c_t$, $f_c = T_c^{-1} c_c$, and $\{f_t^{\text{opt}}, f_c^{\text{opt}}\}$ denotes the set that represents the extracted optimal texture

and cartoon after many iterations. From the Eq. (3), we can get the unknown morphological components rather than the representation vectors. And, there exists a fast numerical algorithm called the block-coordinate relaxation method[12] for Eq. (3). As we know, it is very hard to choose the dictionaries which are perfectly appropriate to represent ϕ_t and ϕ_n the different morphological components. How to choose the proper dictionaries will be discussed in the following Subsection 1.1. Based on the above discussion, the texture and cartoon parts of the image under the complicated background can be decomposed successfully.

1.1 Dictionaries for the Morphological Components

To solve the minimization task of Eq. (3), we must seek two appropriate dictionaries to represent the cartoon and texture parts. We discuss the dictionary for the cartoon part. From the Ref. [11], we can get the conclusion that the curvelet transform (CT) shows good performance for representing the piecewise smooth functions (cartoon parts) sparsely. It deals effectively with local linear singularities in two-dimension.

1.1.1 *Fast Curvelet Transform*

Initial introduction of CT technique was originally introduced by Candes and Donoho in 1999 as a result of the increasing demand of the presence of effective multi-resolution analysis that has the ability to overcome the drawbacks of wavelet analysis. The transform was designed to represent edges and other singularities along curves much more efficiently than traditional transforms (i.e. using significantly fewer coefficients for a given accuracy of reconstruction). Later, based on a frequency partition technique, the same authors proposed a considerably simpler second-generation CT. This second-generation CT is meant to be easier to understand and use. It is also faster and less redundant compared with its first-generation version[13]. The new digital implementations of CT use the unequally spaced Fast Fourier Transform (USFFT), i.e., it uses the "Cartesian coronae" window based on concentric squares (instead of circles) to localize the Fourier transform domain near the sheared wedges obeying the parabolic scaling. The details of the implement can refer to Ref. [13]. Analyzing the CT coefficients, it can be found that the values of CT coefficients are determined by the situation how they are aligned in the real image, the more accurate a curvelet aligned with a given curve in an image is, the higher is its coefficient value. Consequently, by arranging the coefficients of each level and taking the most significant part of them, this will enhance the cartoon part information that represents the important part of the image to us.

1.1.2 *Texture Sparse Representation*

It is difficult to choose a dictionary for a sparse representation of periodic smooth functions (texture part). The texture is shown as a kind of repeated mode with different deformations and different directions in the given different areas of the figure. It is difficult to replace the human visual system for distinguishing the texture only by examining the periodic change of the pixel gray level. So we must imitate the human visual perception mechanism. According to Heisenberg uncertainty principle, Gabor function[14] has the minimum area of time-frequency window and can encode visual signals effectively. In this paper, we will use the two-dimension Gabor function as the atom production function which forms the perceptive dictionary matching the texture part.

1.2 Cartoon Part Extraction Algorithm

The details of the image decomposition algorithm combining Eq. (3), CT, Gabor Transform (GT) are introduced as follows:

Step 1 Initialize the number of iterations m and iteration termination symbol ε, assume initial threshold

$$\lambda^0 = \max\{\|T_{\mathrm{CT}} f\|_\infty, \|T_{\mathrm{GT}} f\|_\infty\},$$

texture part $f_t = 0$ and cartoon part $f_c = 0$.

Step 2 Iterate (m times).

(1) Update the cartoon part. Calculate the residual

$$R^j = f - f_c^{j-1} - f_t^{j-1}$$

assuming the current estimates of f_t^{j-1} is fixed, and the coefficients of CT:

$$b_c = T_{\mathrm{CT}}(R^j + f_c^{j-1}).$$

Estimate the current coefficients $\tilde{b}_c$ by hard thresholding with threshold λ^j. Reconstruct

$$f_c^j = T_{\mathrm{CT}}^{-1}\tilde{b}_c.$$

(2) Update the texture part. Calculate the residual

$$R^j = f - f_c^{j-1} - f_t^{j-1}$$

assuming the current estimates of f_c^{j-1} is fixed, and the coefficients of GT:

$$b_t = T_{\mathrm{GT}}(R^j + f_t^{j-1}).$$

Estimate the current coefficients $\tilde{b}_c$ by hard thresholding with threshold λ^j. Reconstruct

$$f_t^j = T_{\mathrm{GT}}^{-1}\tilde{b}_t.$$

Step 3 Update the threshold. Update the threshold by

$$\lambda^n = \lambda^0 - (n-1)\frac{\lambda^0 - \lambda_{\min}}{m-1},$$

here $\lambda_{\min} = \gamma\sigma$, γ is an integer (in practice, $\gamma = 3$) and σ denotes the standard deviation of the noise.

Step 4 Set the termination condition of the iteration. If $f_n^j - f_n^{j-1} > \varepsilon$ return Step 2, else finish.

After the image decomposition algorithm, we can extract the f_c as the cartoon part of the image.

2 Key Points Descriptor Based on SIFT Algorithm

SIFT provides a comprehensive analysis of several local descriptors with the high performances and relatively low computational cost. Given an image, SIFT features are detected at different scales by using a scale space representation implemented as an image pyramid. The pyramid levels are obtained by Gaussian smoothing and sub-sampling of the image resolution while interest points (key points) are selected as local extrema (min/max) in the scale space. This idea is referred to as the Difference of Gaussians (DoG) approach. Once these key points are detected, SIFT descriptors are computed at their locations in both image plane and scale-space. Each descriptor consists in a histogram of 128 elements, obtained from a 16 pixel × 16 pixel area around the corresponding key point. The contribution of each pixel is obtained by calculating image gradient magnitude and direction in scale-space and the histogram is computed as the local statistics of gradient directions (8 bins) in 4 pixel × 4 pixel patches of the 16 pixel × 16 pixel area.

The SIFT's ability is used to find distinctive key points that are invariant to location, scale and rotation, there are still several false matches because of the background clutter. Then we need the following steps to improve the matching robustness. In particular, given two images I_1 and I_2, we extract the key points and their SIFT descriptors respectively, every descriptor contains 128 elements. The best candidate match for each key point p_i in I_1 is found by identifying its nearest neighbor from the key points in I_2, which is the key point with the minimum Euclidean distance among their descriptors. Due to the high-dimensionality of the feature space in which some descriptors are much more discriminative than others, a global threshold evaluating the distance between the two descriptors does not work well. To increase robustness, we use the ratio between the distance of the closest neighbor to that of the second-closest one, and comparing it with a threshold, i.e., given a key point in I_1 we compute a Euclidean-distance vector $\boldsymbol{D}_{\mathrm{E}} = (d_1, \cdots, d_n)$ that represents the sorted Euclidean-distance with respect to the descriptors in I_2. The matched key point is found only if the following constraint is satisfied: $d_1/d_2 <$ Threshold. Iterating on each key point in I_1, we can obtain the set of matched points robustly.

We do the matching experiment using the "Lena" image based on the SIFT algorithm. The results are shown in Fig. 1. In the preprocess stage, we resize the original image with 0.6 and rotate the image with 21°. Figure 1(a) shows the matching result based on minimal Euclidean-distance without the threshold. We can see that there are several mismatch key points. After setting a threshold as 0.6, we can see the correct matching results in Fig. 1(b). Since we have obtained the corresponding coordinates of the matching key points, the geometry transform matrix can be inferred by using an affine homography. Let the matched point coordinates be, for the two images, $\boldsymbol{p}_i = [x\ y\ 1]^{\mathrm{T}}$ and $\boldsymbol{p}'_i = [x'\ y'\ 1]^{\mathrm{T}}$, respectively, their geometric relationship can be represented by a 3×3 matrix as:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \boldsymbol{H} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ \boldsymbol{0} & \boldsymbol{1} \end{bmatrix}, \tag{4}$$

where $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, and $\boldsymbol{t} = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$. The vector $\boldsymbol{t}$ is the translation while the matrix $\boldsymbol{A}$ is the composition of rotation non-isotropic scaling transformations. In this experiment, the transformation includes only rotation and scaling, matrix $\boldsymbol{A}$ can be decomposed in SVD form as $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^{\mathrm{T}}$, where $\boldsymbol{S} = \mathrm{diag}(s_1, s_2)$ is a scaling, in which s_1 and s_2 represent the (rotated) x and y direction. Then the scale can be estimated as 0.599 87 and rotation angle as 21.069 8°.

(a) The matching result based on minimal Euclidean-distance without the threshold

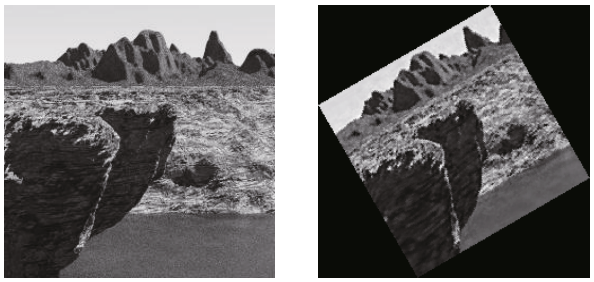
(b) The matching results with a threshold as 0.6

Fig. 1 The matching experiment using the 'Lena' image based on the SIFT algorithm

3 Matching Results Comparison

We compare the registration results between the traditional SIFT algorithm and the proposed method (MCA+SIFT). The original image and the image to be registered are shown in Figs. 2(a) with the size

(256 pixel × 256 pixel) and 2(b). Figure 2(b) is obtained by cropping Fig. 2(a) from the upper left pixel (15, 19) with the size (220 pixel × 220 pixel), scaling with 0.8 and rotation angle with 30°. They are the landscape containing the mountains, meadows, rocks and water. Figures 2(a) and 2(b) conclude the periodic texture and noise parts which are not the stable and distinctive characteristics. Figures 2(c) and 2(d) are the cartoon and residual parts of Fig. 2(a), Fig. 2(c) represents the main geometrical information and Fig. 2(d) represents the textural and noise of Fig. 2(a). From Figs. 2(c) and 2(d), we can see that the different morphological components are separated by the different perceptional dictionaries. This will facilitate to apply the SIFT algorithm to extract the key points and match.



(a) The original image (b) The image to be registered

(c) The cartoon parts of (a) (d) The textural part and noise of (a)

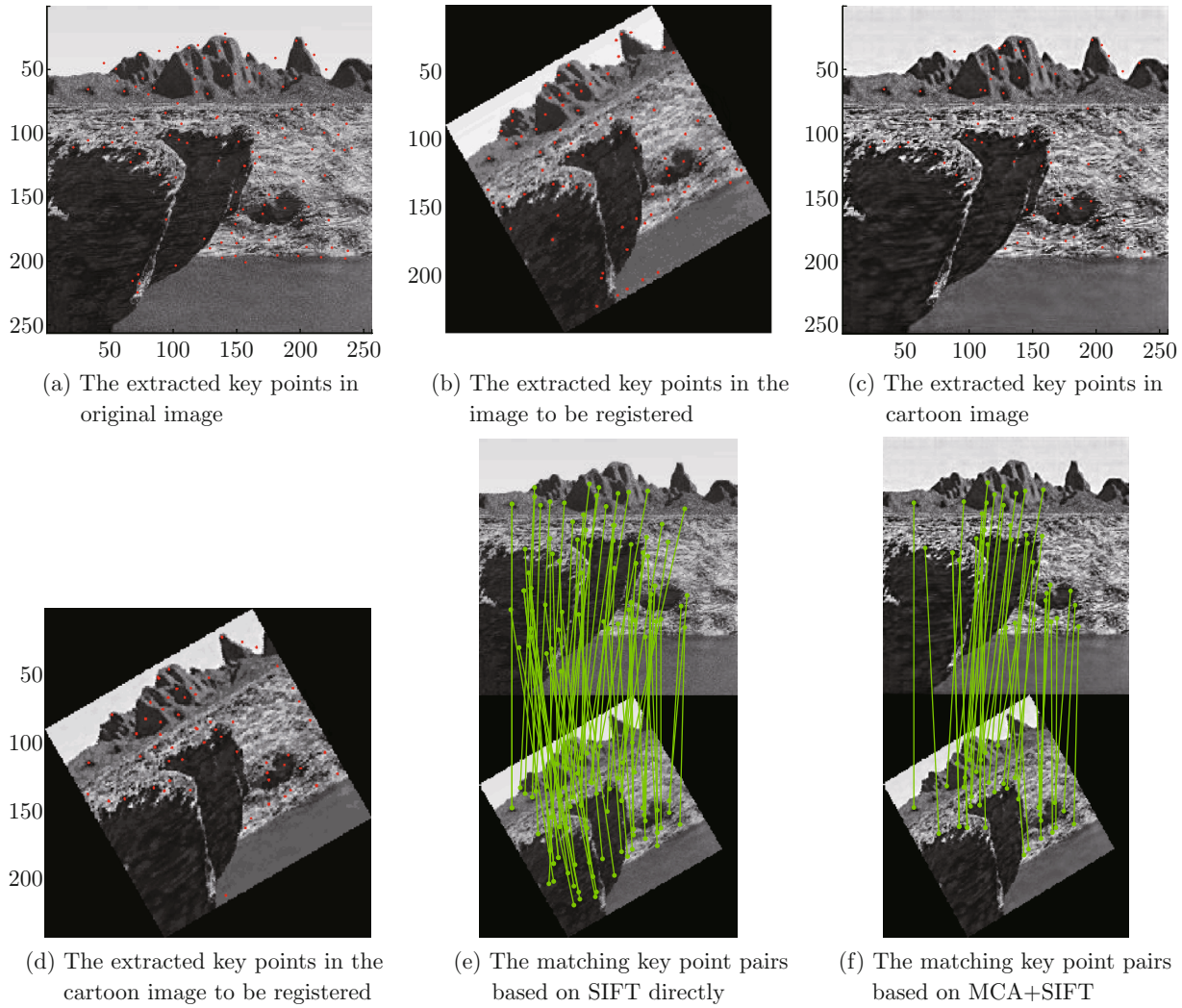
Fig. 2 Separation results based on morphological analysis model

Then we will use the SIFT algorithm to extract the key points by using the Fig. 2(a) as the original image and using Fig. 2(c) as the cartoon part. The images to be registered are Fig. 2(b) and the cartoon part of Fig. 2(b), respectively. Figure 3 is the matching result of the different source images. Figures 3(a), 3(b) and 3(e) are the extracted key points (red point) and matching key point pairs (connected by green line) by using the SIFT algorithm directly for the Figs. 2(a) and 2(b). Figures 3(c), 3(d) and 3(f) are the extracted key points (red point) and matching key point pairs (connected by green line) by using the SIFT algorithm for Fig. 2(c) and the cartoon part of Fig. 2(b). In this experiment, the number of the pyramid levels is chosen as 5, the number of the scale spaces in each pyramid level is 4 and the local extrema threshold is 0.028. To increase matching robustness, we choose the ratio threshold between the distances of the closest neighbor and that of the second-closest one as 0.7. From the results, we can see that the extracted key points from the original image are more than the cartoon part. The number of the key points are 142 and 113 in Figs. 3(a) and 3(b) respectively. The number of the key points in Figs. 3(c) and 3(d) are 89 and 76. Figures 3(a) and 3(b) show that several key points appear in the rock texture and water. Contrast to this phenomenon, Figs. 3(c) and 3(d) show the few key points appear in the texture region of the image. This is due to the main geometrical information inferred from the cartoon part without interference of the unstable texture and noise. This will lead to less computation complexity and a more correct matching. The matching results show that there are several mismatching appearances in the texture region in Fig. 3(e). Based on the Eq. (4) and Fig. 3(e), we can obtain that the translations of x and y directions are $(-16.179, -19.836)$, the scaling is 0.764 5, the rotation angle is 30.075°. Based on the Eq. (4) and Fig. 3(f), we can obtain that the translations of x and y directions are $(-15.839, -19.067)$, the scaling is 0.772 5, the rotation angle is 29.973°. That is to say that all the geometrical transform parameters based on our registration method are more exactly than the ones based on the direct SIFT registration algorithm.

In addition, we use the repeatability measurement to compare the performance of finding matches between the original image registration and the cartoon image registration. The repeatability measurement is computed as a ratio between the number of point-to-point correspondences that can be established for detected points and the mean number of points detected in two images[15]:

$$r_{1,2} = \frac{C(I_1, I_2)}{\text{mean}(m_1, m_2)}, \tag{5}$$

where $C(I_1, I_2)$ denotes the number of corresponding couples, m_1 and m_2 represent the numbers of the detected key points, and mean$(\cdot)$ is the mean function. The following experiment shows the influence of rotation and scaling on the original image registration and the cartoon image registration. The results are shown in Fig. 4(a) which denote the repeatability comparisons between SIFT algorithm and MCA+SIFT algorithm under the different rotation degrees. Figure 4(b) denotes the repeatability comparisons between SIFT algorithm and MCA+SIFT algorithm under the different scaling parameters. From the results, we can see that the MCA+SIFT algorithm can detect the most matches and be stable to rotation and scaling. This indicates that the SIFT can detect more key points than the MCA+SIFT algorithm, which is shown in Fig. 3. However, these key points are not stable, then only a



(a) The extracted key points in original image

(b) The extracted key points in the image to be registered

(c) The extracted key points in cartoon image

(d) The extracted key points in the cartoon image to be registered

(e) The matching key point pairs based on SIFT directly

(f) The matching key point pairs based on MCA+SIFT

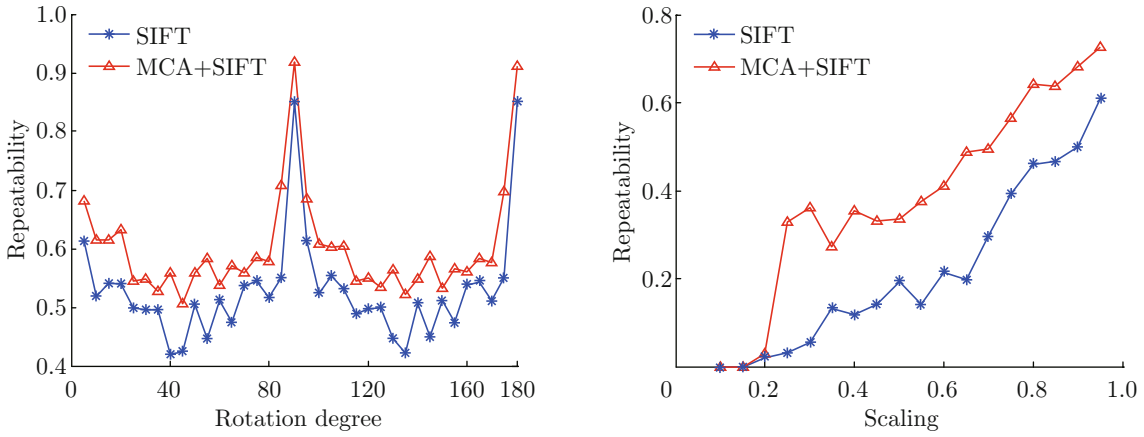
Fig. 3 The matching results of the different registration methods

small portion can be matched according to the constrained Euclidean distance.

In order to examine the registration accuracy, the inferred geometrical transform parameters based on Eq. (4) are computed according to the different rotation angles and scaling. The comparison results between the SIFT and the MCA+SIFT are shown in Fig. 5. The result in Fig. 5(a) shows the inferred rotation degree error with respect to the different rotation degrees of the image to be registered. The blue line with the mark (Δ) indicates the inferred rotation degree error by using the SIFT registration method, the red line with the mark ($*$) indicates the one by MCA+SIFT. From the results, we can see that the inferred rotation degree error based on MCA+SIFT is smaller than SIFT, the mean error based on MCA+SIFT is $-0.018\,2°$ and the mean error based on SIFT is $0.020\,7°$. The comparison result in Fig. 5(b) shows the inferred scaling error with respect to the different scaling of the image to be registered. With the increase in the amplitude of the scaling-down, the inferred scaling errors of the two methods have a tendency to increase. In most cases, the inferred scaling error based on the MCA+SIFT is smaller than the SIFT.

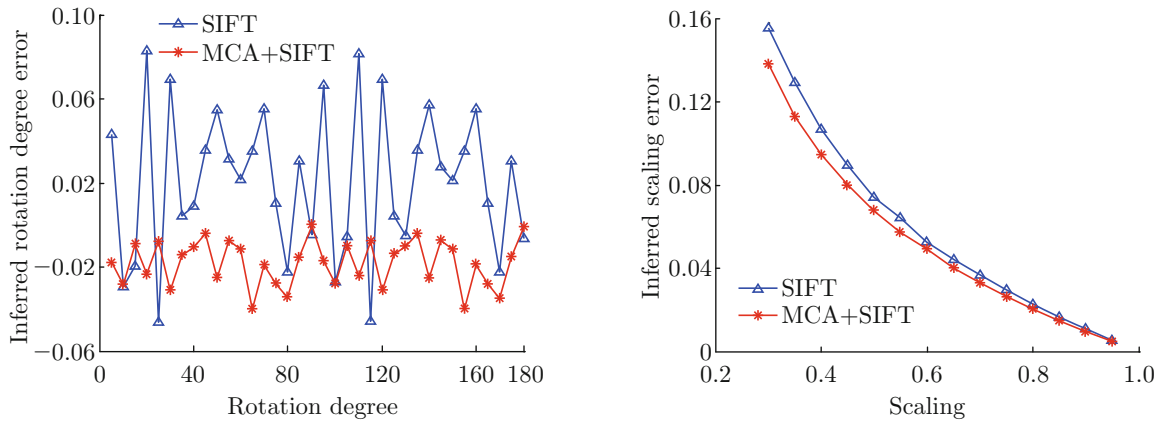
In addition, we compare the registration robustness between the SIFT and MCA+SIFT method. The different level noise is added into Figs. 2(a) and 2(b). The MCA+SIFT method can still separate the cartoon part, the texture part and the noisy residuals from the synthetic image. In contrast, the direct SIFT method needs to use the wavelet de-noising method[16] to remove the influence of the noise. The comparison results between the SIFT and the MCA+SIFT are shown in Fig. 6. The inferred geometrical transform parameters are computed according to the noise with the different standard deviation. Figures 6(a) and 6(b) show the inferred rotation degree error and the inferred scaling error under the condition of the different level noise. Here the stable rotation degree equals to 20° and the scaling parameter equals to 0.5.

The line with the mark (Δ) and the line with the mark ($*$) indicate the inferred rotation degree errors and the inferred scaling errors based on the SIFT and MCA+SIFT registration methods respectively. From the results, we can see that the inferred rotation degree error based on MCA+SIFT is smaller than SIFT, the mean rotation errors based on SIFT and MCA+SIFT are 0.002 7° and 0.000 11°, respectively. The mean scaling errors based on SIFT and MCA+SIFT are 0.076 1 and 0.072 1.



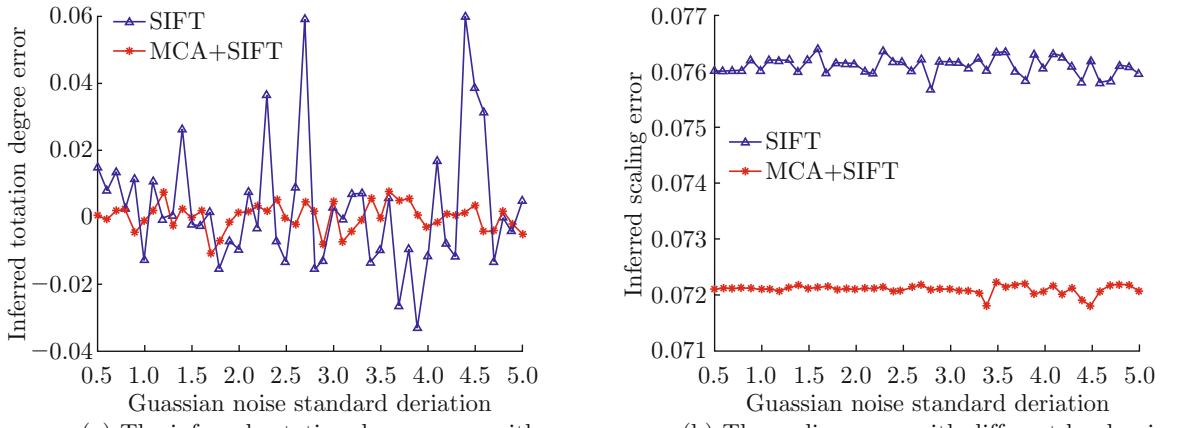
(a) The comparisons for the rotation (b) The comparisons for the scaling

Fig. 4 The influence of rotation and scaling on the SIFT and MCA+SIFT



(a) The comparisons for inferred rotation error (b) The comparisons for inferred scaling

Fig. 5 The registration precision comparion results between the SIFT and the MCA+SIFT



(a) The inferred rotation degree error with different level noise (b) The scaling error with different level noise

Fig. 6 The registration robustness comparion results between the SIFT and the MCA+SIFT

4 Conclusion

This paper has proposed a feature-based image registration method by a fusion of morphological component analysis and SIFT algorithm. This method can solve the mismatching phenomenon under the complicated background, in which SIFT algorithm cannot easily and stably construct the key points. Through constructing the morphological component separation model based on the visual perception mechanism, we can obtain the distinctive and stable key points in the cartoon part of the original image to match the deformed images. Compared our registration method with the direct SIFT registration, a good matching performance, such as correctness and repeatability, can be seen. Also, the matching results can serve as a preprocessing of some further processing such as image fusion and stitching.

References

[1] PENG X, XU J, ZHOU Y, et al. Highly parallel line-based image coding for many cores [J]. *IEEE Transactions on Image Processing*, 2012, **21**(1): 196-206.

[2] YANG J, YU K, GONG Y, et al. Linear spatial pyramid matching using sparse coding for image classification [C]//*Computer Vision and Pattern Recognition, 2009*. Miami, Florida: IEEE, 2009: 1794-1801.

[3] PAL S, MITRA M. Detection of ECG characteristic points using multiresolution wavelet analysis based selective coefficient method [J]. *Measurement*, 2010, **43**(2): 255-261.

[4] MISRA I, MOORTHI S M, DHAR D, et al. An automatic satellite image registration technique based on Harris corner detection and Random Sample Consensus (RANSAC) outlier rejection model [C]// *2012 1st International Conference on Recent Advances in Information Technology (RAIT)*. India: IEEE, 2012: 68-73.

[5] KITCHEN L, ROSENFELD A. Gray-level corner detection [R]. Maryland: Maryland Univ College Park Computer Vision Lab, 1980.

[6] DRESCHLER L, NAGEL H H. Volumetric model and 3D trajectory of a moving car derived from monocular TV-frame sequences of a street scene [J]. *Computer Graphics and Image Processing*, 1982, **20**(3): 199-228.

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

[8] ANDREU F, CASELLES V, DIAZ J I, et al. Some qualitative properties for the total variation flow [J]. *Journal of Functional Analysis*, 2002, **188**(2): 516-547.

[9] VESE L A, OSHER S J. Modeling textures with total variation minimization and oscillating patterns in image processing [J]. *Journal of Scientific Computing*, 2003, **19**(1/2/3): 553-572.

[10] STARCK J L, ELAD M, DONOHO D. Redundant multiscale transforms and their application for morphological component separation [J]. *Advances in Imaging and Electron Physics*, 2004, **132**: 287-348.

[11] STARCK J L, DONOHO D L, CANDÈS E J. Astronomical image representation by the curvelet transform [J]. *Astronomy & Astrophysics*, 2003, **398**(2): 785-800.

[12] BRUCE A G, SARDY S, TSENG P. Block coordinate relaxation methods for nonparamatric signal denoising [C]//*Aerospace/Defense Sensing and Controls*. USA: International Society for Optics and Photonics, 1998: 75-86.

[13] CANDES E, DEMANET L, DONOHO D, et al. Fast discrete curvelet transforms [J]. *Multiscale Modeling & Simulation*, 2006, **5**(3): 861-899.

[14] DAUGMAN J G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters [J]. *Journal of the Optical Society of America A: Optics Image Science and Vision*, 1985, **2**(7): 1160-1169.

[15] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10): 1615-1630.

[16] SARDY S, TSENG P, BRUCE A. Robust wavelet denoising [J]. *IEEE Transactions on Signal Processing*, 2001, **49**(6): 1146-1152.