A Method of Image Registration Based On Best Similarity of Local Geometric Figure

Zetao Jiang
Guilin Uniersity of Electronic Technology, Guilin,China
Nanchang Hangkong University, Nanchang, China
Email: zetaojiang@126.com

Le Zhou and Liwen Zhang
Nanchang Hangkong University, Nanchang, China
Email: {zhoule888,zhangliwen}@126.com

***Abstract*—Image Registration is an important part of computer vision. We propose a method of image registration by obtaining best similarity of local geometric figure that utilizes opposite core difference (OCD) of corresponding local figure. This method gets initial matching after describing precisely SIFT points by constructing feature subspaces based on the detection of SIFT feature points. Then we describe the corresponding similarity by OCD of local figure constructed by SIFT points and choose the feature points that possess highest similarity measure as point set to compute projective transformation matrix T_{opt}. Experiments have proved that the precision of the matrix T_{opt} and the Image matching is at a high level.**

***Index Terms*—SIFT feature subspace, local geometric similarity, projective, OCD**

I. INTRODUCTION

Image matching is one of key technology of computer vision. There are two kinds of methods[1-5] for it: based on feature and based on density. One method based on feature extracts image feature mainly by some algorithms and the another one estimate similarity by image distribution information, the more similar, the higher correct registration is.

The goal of image registration is to get precise transformation from one image to another one. The main procedure is from Coarse to Fine[15,16,19]. Recently, a lot of methods [12-20] of image registration have existed and most of these are focused on feature point such as SIFT [6], Harris [7], SURF [8] points. Harris corner is invariant to rotation, noise and illumination changes of image. But it is not invariant for scale changes. Compared with Harris, SIFT points remedy the drawback of Harris. Thus it's robust to projective transformation of image. And SURF point is the improvement of SIFT. The dimension of SURF descriptor is half of SIFT, but the precision is reduced. Accordingly considering of precision, we choose SIFT points to get initial matching. However, we find that some SIFT corresponding points can't get absolute correspondences even if they are matched in vision. Actually the difference may be about 1-2 pixel. After computing transformation matrix, the image that obtained by the transformation can't coincide with another absolutely, the ghost will appear as showed in Figure 1(a):

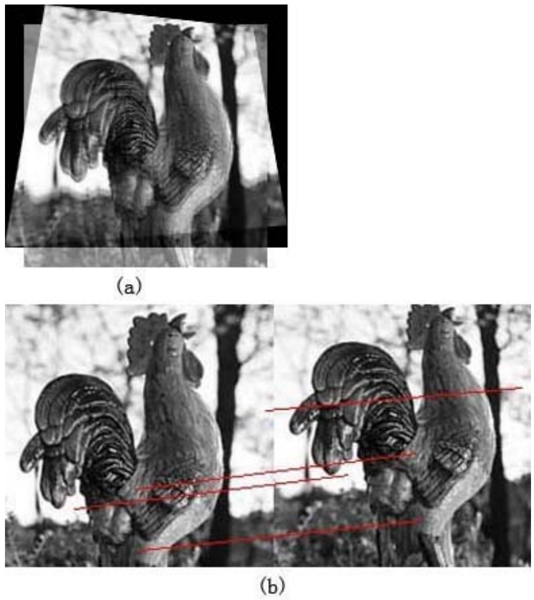

Figure 1(a). The superposition from two images.
(b). Four point pairs used to compute transform matrix

In Figure 1(b), we get transformation by the 4 corresponding points. Then we get the transformed image and correspond to another, as show in Figure 1(a). Hence, this paper will compute OCD of local geometry constructed by 4 corresponding points respectively and we regard OCD as similarity measure. By several iterations, we compute the best transformation matrix T_{opt} through the points that possess the minimum of OCD.

Tal Hassner[9] had studied SIFT algorithm and its scale theory and proposed the concept of SIFT subspace. They constructed basis vectors spanning the scale space of descriptors and got the scale-less SIFT descriptors ultimately. This theory combines the descriptors under different scales, so it ascends the precision of descriptors. We will apply this theory to our feature point descriptor in order to get high precision before initial matching.

doi:10.4304/jsw.9.4.964-969

II. Multi-Scale Feature Detection and Initial Matching

A. SIFT Point Detection

The SIFT algorithm first constructs Gaussian pyramid. For example, we intend to construct O octaves, S levels images and appoint scale value δ for each image defined as below:

$$\delta(o,s) = 2^{o-1} \times k^{s} \times \delta_0$$

Where o and s represent oth and sth image. $k = 2^{1/S}$, δ_0 means initial scale value.

Simply, the scale space $L(x, y, \delta)$ of image $I(x, y)$ will be constructed:

$$L(x, y, \delta) = G(x, y, \delta) * I(x, y)$$

$$G(x, y, \delta) = \frac{1}{2\pi\delta^2} e^{-(x^2+y^2)/2\delta^2}$$

Then, Gaussian DOG images can be obtained by DOG operator:

$$D(x, y, \delta) = [G(x, y, k\delta) - G(x, y, \delta)] * I(x, y)$$

From Gaussian DOG images, we can see the changes of pixels. No changes, no features. The points that change heavily may be features. DOG images portray the contour of object. Feature points are made up of 3-D local extreme points of DOG images. After detecting extreme points we can't regard them as SIFT points until the heavy edge corresponding is eliminated.

B. SIFT Subspace And Feature Descriptor

SIFT descriptor can be obtained by gradient histogram of 16*16 around the feature point. As show in Figure 2, first we compute the main orientation of feature point and then we divide 16*16 into 16 small regions with the size 4*4. Next, we make a statistics about 8 gradient histogram of every 4*4 window and obtain the descriptor vector whose dimension is 4*4*8=128. The descriptor describes the change of gradient orientation around the feature point. Therefore it is robust to delight, scale and rotation. Due to the elimination of edge corresponding, it is also robust to noise.

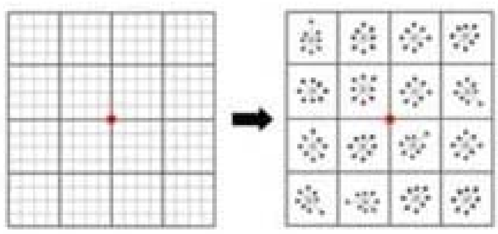

Figure 2. The 16*16 window that computes feature point descriptor

We will expand the descriptors to scale space. For one feature point, we compute the descriptors in different scales and span them into descriptor space called feature subspace. The basis vector of this subspace is used as precise descriptor of the feature point. This shows in Figure 3:

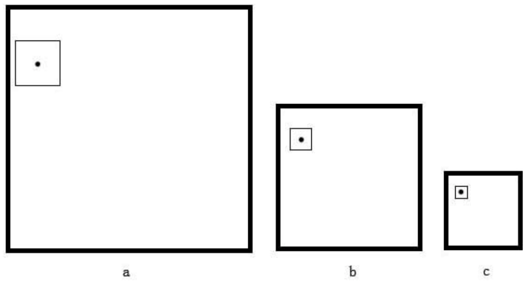

Figure 3. (a),(b) and (c) declare the window around the feature point in different scale

In above figure, a, b and c describe the shape of image in different scale $\delta_a, \delta_b, \delta_c$ for point P. Let $\delta = \{\delta_1, \delta_2, \delta_3, ..., \delta_i\}$ be the scale space of image. Let v_i be the feature vector computed in different scales δ_i, then the feature subspace that derives from v_i is defined as $V = \text{span}\ (v_1, v_2, v_3, ..., v_i)$, and the basis vectors of V is recorded as $v_{i1}, v_{i2}, ..., v_{ik}$. Thereby $\dim(V) = k$. The descriptor of P can be obtained by the linear combination of basis vectors that defined as:

$$Descripter(P) = w_1 v_{i1} + w_2 v_{i2} + ... + w_k v_{ik}.$$

Where w_j is the weight of v_{ij}, and in order to retain unitization, they must satisfy $\sum_{j=1}^{k} w_j = 1$. The choice of w_j depends on the contribution measure of image whose descriptor vector is v_{ij}.

C. Initial Matching Of Feature Points

After the detection and description of features, we utilize Euclidean distance to get initial matching pairs. Let X and Y be SIFT point set to be matched. We get initial matching by steps:

1) For one point x_i of X, we compute the distance $dist(x_i, y_j)$ between x_i and y_i of Y;

2) The minimum and second minimum distance D_m and D_n are chosen to compute the ratio: $ratio = \frac{D_m}{D_n}$;

3) We set a ratio threshold T_{ratio}. If $ratio < T_{ratio}$, x_i and y_i are matched. Otherwise, stop and compute next point x_{i+1}, return step 1);

4) Repeat the above steps until all points of X are computed.

However, there are many wrong pairs in the corresponding points. Many factors account for it such as the same or similar object in an image. In spite of different object, the feature point descriptors may be similar. These outliers emerge mainly without considering the external figure structure of the point. Hence, this paper will take it into consideration by constructing local geometry and getting their similarity. The high similarity geometry is, the more correct the point pair is.

III. LOCAL GEOMETRY SIMILARITY MEASURE

A. Geometry Similarity Measure

By the analysis of last part in Section II, we know the projective transformation matrix T computed by incorrect point pairs is incorrect obviously. In order to get precise matrix T, we choose the most similar local points pairs that defined as:

$$T_{opt} = \arg\max S(T(I_R), I_M)$$

Where $S(T(I_R), I_M)$ represents the similarity of image region I_R transformed and matched image I_M.

Usually, there are a lot of methods to describe the similarity of two image regions. Jeongtae[10] proposed a robust correlation coefficients to describe it. He chose a non-negative density function f_0 and weight function w_0 and computed mean and covariance matrix and achieved the correlation coefficient ρ. Then the best transform T will be obtained through maximum correlation coefficient $\hat{\rho}$. Often the methods that compute correlation coefficients focus on Sample Correlation Coefficient (SCC) and Maximum Likelihood Estimates (MLE).

SCC computes statistics directly by samples, defined as follow:

$$\hat{\rho}(X,Y) = \frac{\hat{C}_S(X,Y)}{\sqrt{\hat{\sigma}_X^2 \hat{\sigma}_Y^2}}$$

Where sample means, sample variances, and sample covariance are defined :

$$\hat{C}(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(X_i - \overline{X})(Y_i - \overline{Y})$$

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$$\overline{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i$$

$$\hat{\sigma}_X^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2$$

$$\hat{\sigma}_Y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \overline{Y})^2.$$

The process of MLE is: 1) Define correlation coefficient of X and Y as ρ whose joint distribution function is $f_{XY}(x,y)$; 2) Compute maximum likelihood estimates according to sample parameter; 3) Get correlation coefficient through maximum likelihood estimates. Usually the result will be reached by lots of iterations.

Compared with MLE, SCC is simple and speedy. While an adequate distribution function should be chosen and many iterations are needed in SCC. However, the correlation coefficient obtained by SCC is not unbiased, even for normal distributions, it is asymptotically unbiased in that case. Therefore, it is not robust to incorrect point pairs. At the same time, it is also time-consuming.

B. Opposite Core Difference (OCD)

Based on above analysis, this paper proposes a method that can obtain best similarity by local geometry. We will describe similarity by OCD of local geometry.

The local regions of two images are shown in Figure 4, let the corresponding points be $P = \{P_1, P_2, P_3, P_4\}$ ($P_i = (X_i, Y_i)$), the cores are shown below:

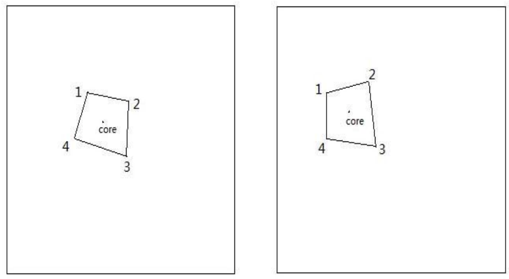

Figure 4. The quadrangles constructed by four points of two images and their cores.
The number in the figure is corresponding index.

Let T_0 be the transformation of P, it means $Y_i = T_0(X_i)$, thereby the cores satisfy $core_Y = T_0(core_X)$.

The property called image core invariance[11]. Based on it, we define the change of cores as similarity measure:

$$S(core_X, core_Y) = f_0(\| T_0(core_X) - core_Y \|)$$

Where $core_X$ and $core_Y$ are cores of two images. f_0 is a non-negative monotonically increasing function. Generally exponential function whose rate of change is obvious is chosen in order to outstand the change value.

Finally, the best transformation is got by minimizing above expression. The pseudo-code of this section described as follow:

// $P: P = \{P_1, P_2, P_3, P_4\}$ is corresponding point
// set and they are not collineation.
// Num: The total number of different choice of P.
// $T_{\text{iteration}}$: Iteration threshold.
$i = 0$;
$Init(S_{\min})$;

```
while(i < min(T_iteration, Num))
{
    T_i = GetTransform(P);
    core_X = GetCore(X);
    core_Y = GetCore(Y);
    S(core_X, core_Y) = f_0(|| T_0(core_X) - core_Y ||);
    if(S(core_X, core_Y) < S_min)
    {
        S = S(core_X, core_Y);
        T_opt = T_i;
    }
    i++;
    Update(P);
}
```

After all, the concrete steps of this paper are:

(1) Feature Detection. The feature point sets P_1 and P_2 of image I_R and I_M are obtained from SIFT algorithm;

(2) Multi-Scale Feature Description. Apply SIFT subspace theory to every SIFT point $p_{1i}(p_{1i} \in P_1)$ and $p_{2i}(p_{2i} \in P_2)$ to get precise descriptors.

(3) Initial Matching. Get initial matching by Section2.3 and initial point pairs $P(P=(P_R, P_M))$;

(4) Get Points That Possess Highest Similarity. Apply Section3.2 to 4 points that possess the highest similarity;

(5) Compute Transformation Matrix T_{opt} by these 4 points.

IV. Experiments and Verification

The source images of the paper show as follow:

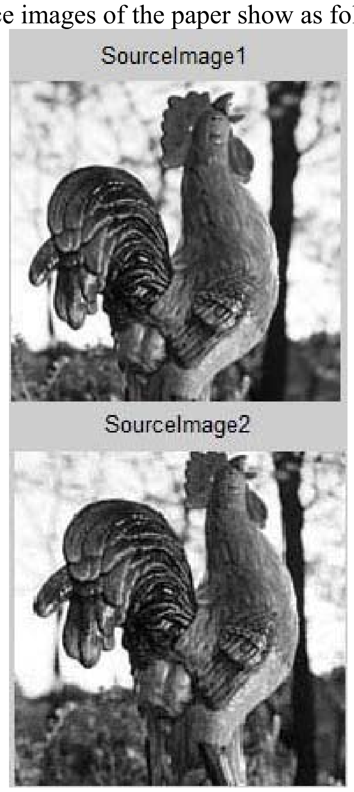

Figure 5. The source images

SourceImage1 and SourceImage2 are the same scene that is got from different view.

We apply Matlab to prove that. In the experiment, after first iteration, the four corresponding points show in Figure 6:

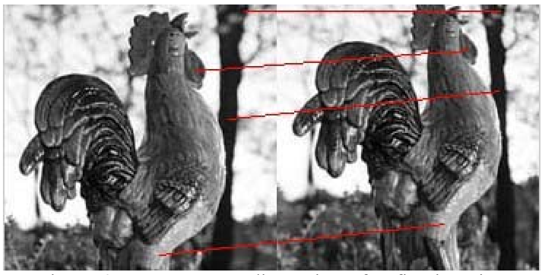

Figure 6. Four corresponding points after first iteration

From TABLE I, the Euclidean distance is 7.2596 after first iteration. Although they are matched in vision, but the difference exists between the cores. The incorrect matching of one or more points of 4 points can account for it. So more iterations are needed to obtain best transformation T_{opt}. The number of iteration and their distance are:

The projective matrix when the OCD is 0.3173:

$$T_{opt} = \begin{bmatrix} 1.1100 & 0.0452 & 7.6545e\text{-}04 \\ -0.1150 & 0.9996 & -9.0902e\text{-}04 \\ 12.4259 & -32.9569 & 1.0362 \end{bmatrix}$$

Therefore, the final superposition image is shown in Figure 7:

Figure 7.The superposition image between the first image transformed and the second image

Compared with Figure 1(a), the matching degree is enhanced conspicuously. However, the part of cock' header is smooth in vision. The interpolation algorithm of Matlab may account for it.

Besides, another experiment results have verified the validity of our algorithm in Figure 8.

Figure 8.The top two images are source images. The superposition result is shown on the bottom.

TABLE I.
THE ITERATION NUMBERS AND THEIR OCD(EUCLIDEAN DISTANCE)

Iteration number	OCD(Euclidean distance)
1	7.2596
2	4.2156
4	2.2603
6	0.7736
8	0.3173

V. CONCLUSION

This paper proposes a simple and practical method that computes the transformation matrix between images. Core invariance is applied in image matching. In order to get precise descriptor, we get feature descriptor in different scales. In this paper, we chose 4 corresponding points at random and have not studied how to choose geometry points to improve speed. In theory, the quadrilateral is neither too large nor too small and more explorations will be taken.

ACKNOWLEDGEMENTS

This paper is sponsored by Nature Science Foundation of China (61272216) and innovation Foundation of aerospace in China (CASC201102) and Nature Science Foundation of Jiangxi Province(20114BAB201035) and Science technology cooperation of Jiangxi Province(2010EAB01700).

REFERENCES

[1] Hyung Il Koo, "Feature-based Image Registration Algorithm for Image Stitching Applications on Mobile Devices," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.
[2] Chunlei Huo,Chunhong Pan. "Multilevel SIFT Matching for Large-Size VHR Image Registration," *IEEE Geoscience and Remote Sensing Letters*. vol.9,No.2,2012.3.
[3] Y. Bastanlar,A. Temizel. "Improved SIFT matching for image pairs with scale difference," *Electronics Letters*, vol.46, No.5, March 2010.
[4] Mahdi Amiri. RASIM:A "Novel Rotation and Scale Invariant Matching of Local Image Interest Points," *IEEE Transactions on Image Processing*. vol.20, No.12, December 2011.
[5] Pengfei Du. "Improved SIFT Matching Algorithm for 3D Reconstraction from Endoscopic Images,". *Association for Computing Machinery*, 978-1-4503-1060-4/11/0012.
[6] Lowe D G. "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol.2, pp. 91-110, 2004.
[7] Harris C, Stephens M. "A combined corner and edge detector," *Alvey vision conference*, 1988.
[8] Bay H, Tuytelaars T, "Van Gool L. Surf: Speeded up robust features," *Computer Vision–ECCV 2006*. Springer Berlin Heidelberg, 2006.
[9] Tal Hassner, Viki Mayzels. "On SIFTs and their Scales," *Computer Vision and Pattern Recognition*, 2012.
[10] Jeongtae Kim, Jeffrey A. "Intensity-Based Image Registration Using Robust Correlation Coefficients," *IEEE Transaction on Medical Image*. vol.23, No.11, November 2004.
[11] Jianhua Jia, Lichen Jiao. "Affine Invariant Texture Feature Extraction Based on Invariant Centroid,". *Chinese Journal of Electronics*, vol.36, No.10, pp. 1910-1915, 2008.
[12] Shanhu Wang, Hongjian You. "BFSIFT: A Novel Method to Find Feature Matches for SAR Image Registration," *IEEE Geoscience and Remote Sensing Letters*. vol.9, NO.4, July, 2012.
[13] Chunlei Huo, Chunhong Pan. "Multilevel SIFT Matching for Large-Size VHR Image Registration," *IEEE Geoscience and Remote Sensing Letters*.vol.9, NO.2, March, 2012.
[14] Chien-Chang Chen,Zeng-Chuang Jiun. "An Adaptive Reversible Image Watermarking Scheme Based on Integer Wavelet Coefficients". *Journal of Computers*, vol 8, NO.7, pp. 1677-1682, 2013.
[15] Mohammad Izadi, Parvaneh Saeedi. "Robust Weighted Graph Transformation Matching for Rigid and Nonrigid Image Registration," *IEEE Transaction on Image Processing*. vol.21, No.10, 2012.
[16] Zhaoxia Liu, Jubai An. "A Simple and Robust Feature Point Matching Algorithm Based on Restricted Spatial Order Constraints for Aerial Image Registration,". *IEEE Transaction on Geoscience and Remote Sensing*. vol.50, No.2, 2012.
[17] Yinan Yu, Kaiqi Huang. "A Novel Algorithm for View and Illumination Invariant Image Matching," *IEEE Transaction on Image Processing*.vol.21, No.1, January, 2012.
[18] Yudong Zhang,Shuihua Wang. "Muti-channel diffusion Tensor registration via adaptive chaotic PSO,". *Journal of Computers*. vol 6, No.4, pp. 825-829, April, 2011.
[19] Peng Wang, Zhiguo Qu. "A Coarse-to-Fine Matching Algorithm for FLIR and Optical Satellite Image Registration,". *IEEE Geoscience and Remote Sensing Letters*. vol.9, No.4, July, 2012.
[20] Yanxia Wang,Yan Ma. "A Method of Line Matching Based on Feature Points," *Journal of Software*,vol 7, No.7, pp. 1539-1545, 2012.

Zetao Jiang Male, Born in 1961 in Jiangxi Province, professor, doctor. He is adjunct professor of Guilin Uniersity of Electronic Technology, PhD supervisor, academic leader in Guangxi province. He received the B.S. degree in Beijing Normal University, Beijing, China, in 1986. He received the M.S. degree in Tongji University, Shanghai, China, in 1995 and received the Doctor of Philosophy in Northwestern Polytechnical University, Xi'an, China, in 2006. Main activities were research in the field of image processing and computer vision.

He presided key technology research project "Control system based on multi-level security agent", which was awarded the second class prize in the 2009 annual science and technology progress of Jiangxi province. He presided the research of network and information security defense technology based on digital watermarking, which was awarded the first class prize in science and technology achievements of Jiangxi province. He presided and finished Nature Science Foundation of China(60673055), two subject Foundations of province, an open foundation of measurement and control center of province and more than 10 other subjects. Currently, he is presiding Nature Science Foundation of China(60673055) and Nature Science Foundation of Province, which is in the field of 3D reconstruction.

Zhou Le Male, Born in 1988 in Jiangsu Province, and received the B.S degree in Yancheng Teachers' College, Yancheng,China. Currently, he is pursuing the M.S degree in Nanchang Hangkong University, and the main research area is image processing and computer vision.

Zhang Liwen Female, Born in 1989 in Jiangxi Province, and had received the B.S degree in Nanchang Hangkong University. Currently, she is also pursuing the M.S degree in Nanchang Hangkong University, and her research area is image processing and computer vision.