

# Image Matching Algorithm on Wavelet-Edge-SIFT

Sun Xun<sup>1,a</sup> Liu Shun-hua<sup>1,b</sup> Luo Zhi-rong<sup>2,c</sup> Li Wei<sup>3,d</sup>

<sup>1</sup> Graduate Management Department of Institute of NBC Defence, Beijing, 102205, China;

<sup>2</sup> 4th Department of Institute of NBC Defence, Beijing, 102205, China

<sup>3</sup> Training Department of Institute of NBC Defence, Beijing, 102205, China

<sup>a</sup>email: sunxun2373@163.com, <sup>b</sup>email: liu\_shun\_hua@sohu.com, <sup>c</sup>email: lzh760519@sina.com,

<sup>d</sup>email: liwei1128@126.com

**Keywords:** wavelet transform; edge detection; SIFT; WES; Image matching

**Abstract.** Aiming at excessive feature points to be extracted in the SIFT algorithm widely applied in image matching, the paper proposes the following solution strategy: obtaining image edge by wavelet transform, subsequently, applying SIFT algorithm for image matching on the basis of the detected image edge. The result shows that the image edge detection matching WES(Wavelet-Edge-SIFT) algorithm by wavelet transform greatly reduces the calculation amount, it may reduce mismatch, increase matching speed and greatly improve the real-time of algorithm.

## 1 Introduction

At present, the image matching algorithm based on SIFT(scale invariant feature transform) has the following shortcomings when matching an visible image with many similar areas: single matching restraint conditions, hard to efficiently reject mismatch points, high mismatch rate; too large extraction quantity of SIFT feature points, too large statistic scope in the gradient direction and too large scope for establishing feature vectors. These shortcomings cause too much time to be consumed in the SIFT statistic gradient direction and for establishing a descriptive operator[1]. For the solution of such a defect, the paper proposes an edge detection algorithm based on wavelet analysis, i.e. WES(Wavelet-Edge-SIFT) wavelet transform image edge detection matching algorithm.

Edge may reflect the partial mutation information in a scene and it is a quite stable image feature[2].

As a kind of effective mathematical tool quickly developed in the recent years, wavelet transform has been widely applied in the field of image processing. It has the ability of detecting partial mutation and is a kind of effective tool detecting image edge.

In the algorithm, firstly the edge information of a basic referential image is detected by wavelet analysis algorithm, then the edge information of the real-time image is detected, finally, SIFT algorithm is applied to match the real-time image and the basic referential image by the edge information, so to realize precise and fast matching of images.

## 2 WES(Wavelet-Edge-SIFT) wavelet transform image edge SIFT matching algorithm

### 2.1 Image edge detection based on wavelet transform

#### 2.1.1 Brief introduction on the wavelet transform algorithm[3]

The main function of the algorithm is to detect the image edge by wavelet transform. An image edge, from the viewpoint of pixel point, shall be the mutation point of gray value. Just by utilizing the property, the wavelet edge with adjustable direction extracts the edge of an image.

Suppose  $\theta(x, y)$  is a properly smooth two-dimensional function, and after a size parameter  $a$

has been introduced,  $\theta_a(x, y) = \frac{1}{a^2} \theta\left(\frac{x}{a}, \frac{y}{a}\right)$ , the two-dimensional wavelet under the size parameter

$a$  is defined as follows:

$$\begin{cases} \psi^x(x, y) = \frac{\partial \theta(x, y)}{\partial x} \\ \psi^y(x, y) = \frac{\partial \theta(x, y)}{\partial y} \end{cases} \quad (1)$$

The two-dimensional wavelet transform is expressed as follows in the manner of vector:

$$\begin{aligned} a \left[ \frac{\partial(f * \theta_a)(x, y)}{\partial x} \bar{i} + \frac{\partial(f * \theta_a)(x, y)}{\partial y} \bar{j} \right] &= a \left[ \left( f * \frac{\partial \theta_a}{\partial x} \right)(x, y) \bar{i} + \left( f * \frac{\partial \theta_a}{\partial y} \right)(x, y) \bar{j} \right] \\ &= af(x, y) * [\psi^x(x, y) \bar{i} + \psi^y(x, y) \bar{j}] = a[(f * \psi^x)(x, y) \bar{i} + (f * \psi^y)(x, y) \bar{j}] = W_a^x f(x, y) \bar{i} + W_a^y f(x, y) \bar{j} \\ &= a \nabla(f * \theta_a)(x, y) \end{aligned} \quad (2)$$

The module value of the gradient vector is  $Mod[W_a f(x, y)] = \sqrt{|W_a^x f(x, y)|^2 + |W_a^y f(x, y)|^2}$ , the included angle between the gradient vector and the direction of horizontal x axle is  $\alpha = Arg[W_a f(x, y)] = arctg \left[ \frac{W_a^y f(x, y)}{W_a^x f(x, y)} \right]$ . It may be known from the previous analysis that the edge corresponds to the location at the extreme value of  $Mod[W_a f(x, y)]$ . By changing  $\alpha$ , the multi-size extraction of edge may be realized.

The wavelet transform in an arbitrary direction is defined as follows:

$$W_a^\theta f(x, y) = f(x, y) * g(x, y, \theta) = f(x, y) * [\psi^x(x, y) \cos \theta + \psi^y(x, y) \sin \theta] \quad (3)$$

$$\begin{aligned} &= W_a^x f(x, y) \cdot \cos \theta + W_a^y f(x, y) \cdot \sin \theta = \|W_a f(x, y)\| \cdot \left( \frac{W_a^x f(x, y)}{\|W_a f(x, y)\|} \cdot \cos \theta + \frac{W_a^y f(x, y)}{\|W_a f(x, y)\|} \cdot \sin \theta \right) \\ &= \|W_a f(x, y)\| \cdot (\cos(Arg[W_a f(x, y)]) \cdot \cos \theta + \sin(Arg[W_a f(x, y)]) \cdot \sin \theta) \\ &= \|W_a f(x, y)\| \cdot (\cos \alpha \cdot \cos \theta + \sin \alpha \cdot \sin \theta) = \|W_a f(x, y)\| \cdot \cos(\alpha - \theta), \text{ therefore} \\ &Mod[W_a f(x, y)] = \max[W_a^\theta f(x, y)] \end{aligned} \quad (4)$$

Which shows that the square of a directional wavelet transform value is proportional to the square of the gradient module value after the function is smoothened. Therefore, instead of obtaining the gradient module value, we may obtain the directional wavelet transform value. Because of the relation between the gradient value and the wavelet transform value of function, we may seek for the maximum value along the direction of wavelet transform, it is not necessary to calculate the gradient direction and the gradient module value.

The 2-D rotary symmetrical Gaussian Function with variance  $\sigma^2$  is

$$\theta_a(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (5)$$

The two directional derivatives of Gaussian Function are

$$\begin{cases} \theta_a^x(x, y) = -\frac{x}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \\ \theta_a^y(x, y) = -\frac{y}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \end{cases} \quad (6)$$

If the size is  $2^j$ , the 2-D dyadic wavelet transforms in the direction of x and y are

$$\begin{cases} \psi_{2^j}^x(x, y) = -\frac{1}{4^j} \theta^x\left(\frac{x}{2^j}, \frac{y}{2^j}\right) \\ \psi_{2^j}^y(x, y) = -\frac{1}{4^j} \theta^y\left(\frac{x}{2^j}, \frac{y}{2^j}\right) \end{cases} \quad (7)$$

$$\begin{cases} W_a^x f(x, y) = f(x, y) * \psi_a^x(x, y) \\ W_a^y f(x, y) = f(x, y) * \psi_a^y(x, y) \end{cases} \quad (8)$$

$$W_{2^j}^x f(x, y) = f(x, y) * \frac{1}{4^j} \theta^x \left( \frac{x}{2^j}, \frac{y}{2^j} \right) = \frac{1}{4^j} \left( -\frac{1}{2\pi\sigma^4} \right) f(x, y) * x \cdot \exp \left( -\frac{x^2 + y^2}{4^j 2\sigma^2} \right) \quad (9)$$

Let  $K = -4^j 2\sigma^2$ , then the formula (9) is

$$\begin{aligned} W_{2^j}^x f(x, y) &= \frac{1}{K\pi\sigma^2} \cdot f(x, y) * x \cdot \exp \left( \frac{x^2 + y^2}{K} \right) \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(x-m, y-n) \cdot m \cdot \exp \left( \frac{m^2 + n^2}{K} \right) / K\pi\sigma^2 \end{aligned} \quad (10)$$

However, due to the influence of image noise, the correlation of neighborhoods is relatively low, so that a directional filter is hard to accurately position the boundary state of a pixel point (x,y). Owing to the arbitrariness of the direction of the image boundary, the wavelet transform responses in many directions are calculated to search the image boundaries in all possible directions, the transform window is 5×5, as shown in Figure 1. The final edge detection result is the even value of the filtering images in many directions.

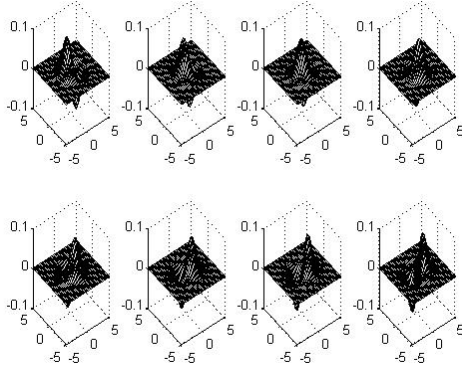


Fig.1. Adjustable wavelet in 8 directions



Fig.2. Gray image A

### 2.1.2 Test on detection of image edge based on wavelet transform

In order to increase the matching speed, the original referential image as input is converted into the format of gray image from RGB format, as shown in Figure 2; the  $A_e$  as the edge of referential image obtained by wavelet transform is shown in Figure 3.

The real-time image B is shown in Figure 4; the referential image  $B_e$  obtained by wavelet transform is shown in Figure 5.

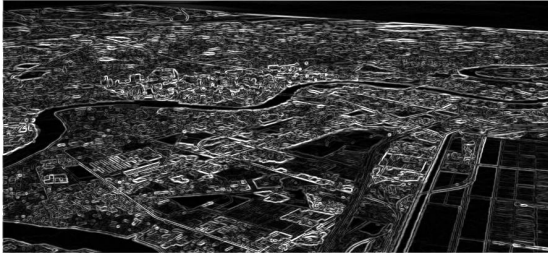


Fig.3.  $A_e$  as the edge of referential image

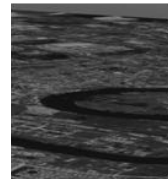


Fig.4. real-time image B

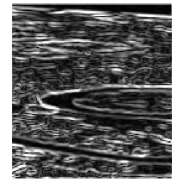


Fig.5.  $B_e$

## 2.2 Steps of SIFT algorithm

Main thought of SIFT algorithm: seeking an extreme point in the size space; then filtering it and finding out a stable feature point; finally, extracting the partial properties of an image around each stable feature point, so to form a partial descriptor and apply it in the subsequent matching. In the invariance contrast experiment made by Mikolajczyk for 10 kinds of partial descriptors including SIFT operator, SIFT and its expanded algorithm were verified to own the most powerful robustness[4] among the similar descriptors, they have been successfully applied in such fields as image registration[5].

SIFT algorithm mainly includes the following steps:[1]

The size space of a piece of image  $I(x,y)$  is defined as  $L(x,y,\sigma)$ , which may be expressed as:

$$L(x,y,\sigma) = G(x,y,\sigma) \otimes I(x,y) \quad (11)$$

In which,  $G(x,y,\sigma)$ ——Gaussian Function with different sizes;  $\otimes$  expresses convolution;  $(x,y)$  is the coordinates of a pixel;  $\sigma$  is the factor of size space.

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (12)$$

A group of continuous Gaussian convolution kernels are utilized to convolve with the original images, so to generate a series of images on size space, the difference between the neighboring size images is *DOG* (Difference of Gaussian) image  $D(x,y,\sigma)$ , i.e.:

$$\begin{aligned} D(x,y,\sigma) &= (G(x,y,k\sigma) - G(x,y,\sigma)) \otimes I(x,y) \\ &= L(x,y,k\sigma) - L(x,y,\sigma) \end{aligned} \quad (13)$$

Each point in *DOG* size space is compared one by one with 26 neighbourhoods including the neighboring locations of the size and the corresponding locations of the neighboring sizes, the attained location and size of the partial extreme value are just the location of the candidate feature point and its corresponding size. Finally, by filtering the feature points with low contrast and located at the edge, a stable feature point may be attained.

The gradient module and direction of a feature point are

$$m(x,y) = \sqrt{(L(x+1,y)-L(x-1,y))^2 + (L(x,y+1)-L(x,y-1))^2} \quad (14)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1)-L(x,y-1))/(L(x+1,y)-L(x-1,y))) \quad (15)$$

A feature point may be assigned with many directions, which may strengthen the matching robustness.

With a feature point as the center, a neighbourhood with  $8 \times 8$  is taken as the sampling window, the formula (14) and (15) are utilized to obtain the gradient value and gradient direction of each pixel, the descriptor of each feature point is a 128-dimensional vector, finally, the descriptor vector of the feature point is uniformized.

When the SIFT feature vector is generated, Euclidean distance is taken as the distance function of SIFT descriptor, the distance-ratio criterion is complied with for matching, i.e. if a feature point has been reached or not by matching is to be determined by the ratio of the nearest neighboring distance and the secondary nearest neighboring distance. When the ratio is less than some threshold value, it is deemed that the matching is successful.

### 3 Test

Firstly, the aircraft are utilized to shoot the scene images in the flying area, these images are jointed and formed an original referential image A; thereafter, the information of the aircraft, such as location(longitude, latitude, height), orientation(pitch angle, yaw angle and roll angle) and time, are utilized to select scene adaptation area C in the basic referential image A and initially position the terrain. The selection of the scene adaptation area is a key step in terrain matching, it directly decides the real-time of terrain matching; finally, the real-time shot image B is matched with the scene adaptation area C by SIFT algorithm, so to realize precise positioning. In the wavelet edge SIFT algorithm, wavelet transform is firstly utilized to obtain the edge  $A_e$  of the basic referential image A, subsequently, the wavelet transform is utilized to calculate the edge  $B_e$  of the real-time image, thereafter the image  $C_e$  of the matching area is calculated,  $C_e = C$ . Finally, on the basis of the real-time shot image  $B_e$  and the scene adaptation area  $C_e$ , SIFT algorithm matching is conducted for the real-time of precise positioning, as shown in Figure 6.

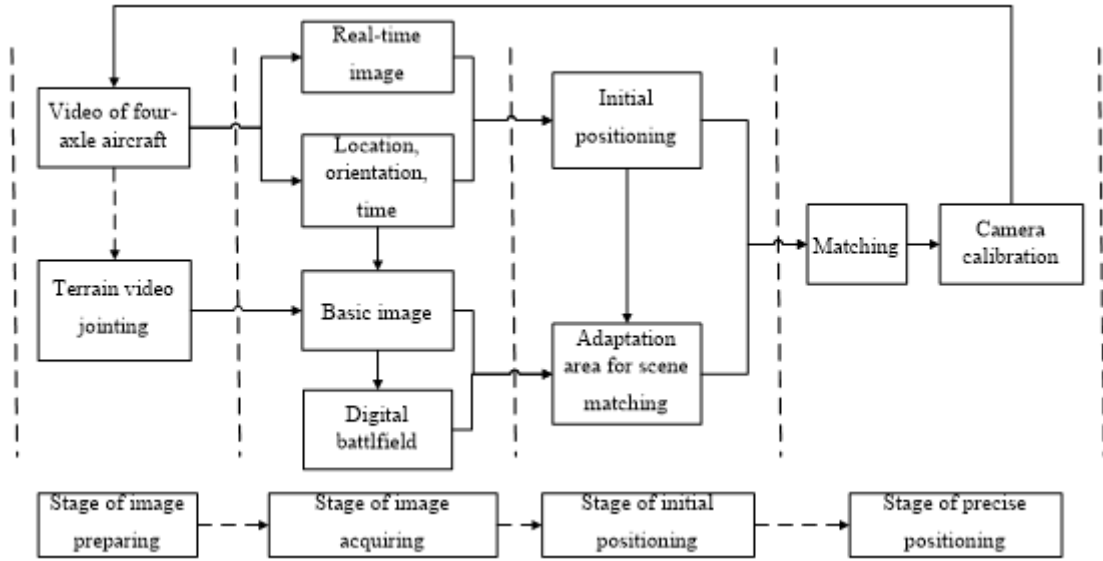


Fig. 6. Working process of image matching

By utilizing SIFT algorithm, the matching process of the original image is shown as Figure 7; By utilizing SIFT algorithm and edge detection, the images are matched, the process is shown in Figure 8; the result after integration is shown in Figure 9.

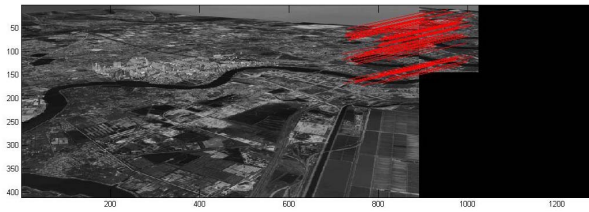


Fig.7. Matching process

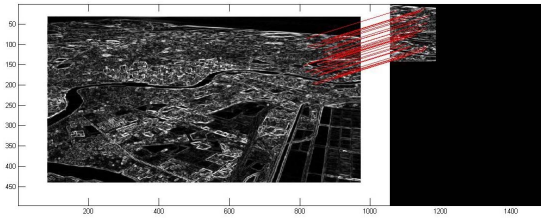


Fig.8. Image edge matching process

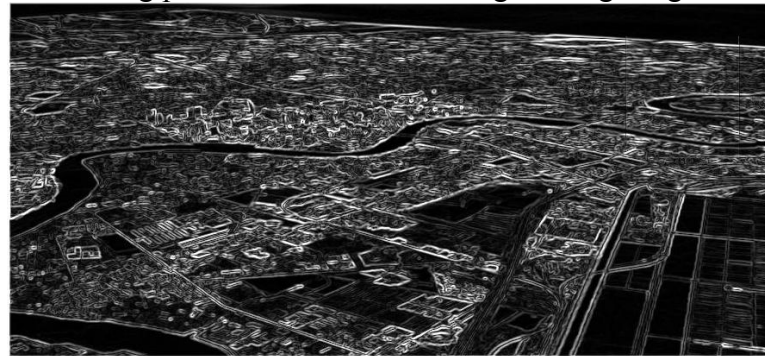


Fig.9. Integrated result by image edge matching

#### 4 Assessment indices and result analysis

By program calculation, the time for extracting feature points on the basis of both algorithms were recorded. The calculation on the edge  $A_e$  of image A was conducted in advance, therefore, the time is not contained into the matching time; the time for calculating the matching area is consistent, for the sake of simplification, such time is also not contained.  $t_1$  is the SIFT algorithm matching time of the real-time image B and the image C of the matching area in the basic referential image A;  $t_2$  is the sum of the time  $t_3$  for calculating the real-time image edge  $B_e$  and the SIFT algorithm matching time  $t_4$  of the image  $B_e$  and the matching area image  $C_e$  in the referential image edge  $A_e$ .  $t_5$  is the time saved by the total time of WES operator than SIFT operator;  $t_6$  is the time saved by the edge detection SIFT matching time and the SIFT matching time of common picture.

Table 1. Consuming time of SIFT and WES

Operator	SIFT	WES	
Time	t1 = 2.882787	t3=0.249279	t4 = 1.918001
		t2 = 2.26728	
Saved time	t5 = t1- t2 = 0.615507 t6 = t1- t4 = 0.964786		

By comparing the matching process in Figure 7 and Figure 9, it is clear that the matching process in Figure 7 is more complex than that in Figure 9, the main reason is overlarge statistic scope in the direction of gradient and overlarge scope for establishing a feature vector caused by excessive extraction amount of SIFT feature points. In addition, it may be seen from Table 1 that, the total time of WES operator is saved than the time of SIFT by  $t_5 = t_1 - t_2 = 0.615507$ S, the SIFT matching time for edge detection is saved than the SIFT matching time for common picture by  $t_6 = t_1 - t_4 = 0.964786$ S. The total time for WES to complete feature matching is less than that of SIFT, on the aspects of matching process and algorithm speed, WES algorithm has apparent advantages.

## 5 Conclusion

The former SIFT algorithm is innovated, improved and integrated, the edge detection algorithm applying wavelet transform is introduced. The algorithm greatly reduces the search scope in SIFT matching process, strengthens operation restraint and avoids misconvergence of operation, etc., the calculation amount of SIFT algorithm is greatly reduced. The experimental result shows that, WES operator greatly increase the real-time of algorithm, it has powerful robustness for linear deformation, brightness variation and noise change of image, in comparison with SIFT algorithm, a higher real-time requirement may be met.

## References:

- [1] Zhao Mengmeng, Cao Jianqiu. SIFT Registration Algorithm Based on Edge Corner [J], Journal of Chongqing Jiaotong University (Natural Science page), Vol. 32, No. 4, Aug 2013
- [2] YangZeqian, Liu Guangqi, Wu Yongguo, Li Deqiang. Image Edge Detection Matching Algorithm Based on Wavelet Transform[J], Microcomputer Information (Automatization of Measurement & Control), 2010, Vol.26, No. 4-1.
- [3] Yang Fusheng, Engineering Analysis and Application of Wavelet Transform, Science Press, 1999.
- [4] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.
- [5] LI Ling-ling, LI Cui-hua, ZENG Xiao-ming, et al. An automatic image registration method based on SIFT and Harris-Affine features matching[J]. Huazhong Univ. of Sci.& Tech: Natural Science Edition, 2008, 8(36):13-16.