# A Fast Method for Image Matching and Registration Based on SIFT Algorithm and Image Pyramid

**Chuang Guo[1], Fangxiu Jia[1*] and Wei Tang[1] and Pan Huang[1]**

[1]ZNDY Ministerial Key Laboratory, Nanjing University of Science and Technology, Nanjing, Jiangsu, 210094, China

*Corresponding author's e-mail: jiafangxiu@126.com

**Abstract.** In this paper, a fast image registration algorithm based on SIFT is proposed, which is slow in extracting feature points from high-resolution images. In this paper, the original image is transformed into an image pyramid to obtain a sub-sampled image with a lower resolution. At the same time, the Laplace image pyramid is obtained. Feature points of the sub-sampled image are extracted through SIFT algorithm, and then they are fused with the original image to obtain feature points of the original image. Experimental results show that the algorithm can quickly find the feature points and obviously improve the computational efficiency.

## 1. INTRODUCTION

With the development of image processing technology in recent years, image registration has been widely used in medical, machine vision, military and other fields [1]. Image registration refers to the process of stitching and fusing two or more images acquired by different sensors under different external conditions (different shooting angles, different shooting times, different distortion degrees, different transformation conditions, etc.) [2]. Currently, image registration methods are mainly divided into region-based image registration methods and feature-based registration methods. Region-based registration method is mainly based on the alignment of the correlation intensity between image pixels. It is characterized by the use of as much data as possible in the image. The registration accuracy is very high, but the initialization of registration parameters is relatively complicated [3]. The feature-based registration method is to extract the feature space of the image, such as edge, curve, corner, high curvature point, plane, etc., as the feature to estimate the transformation matrix of the image [4].

With the rise of machine vision, image processing technology has become an indispensable part of machine vision. At the same time, the image contains a large amount of information, so the image processing algorithm is required to improve the processing speed. Since the traditional image registration algorithm cannot meet the requirements of real-time, high efficiency and high precision, experts have put forward a lot of excellent registration algorithms. With the rapid development of feature-based image registration methods. Lowe proposed the Scale Invariant Features Transform (SIFT) algorithm for the first time at the international conference on computer vision in 1999 and improved it in 2004 [5]. SIFT algorithm has a good effect in image translation, rotation, scaling, Angle transformation and light variation, so it is widely valued [6]. The SIFT algorithm has the characteristics of scale and rotation invariance, strong anti-noise performance, etc., but SIFT algorithm calculation speed is slow, large amount of data, high complexity, long time consumption [7]. The Speed-Up Robust Features (SURF) algorithm was proposed on the basis of SIFT algorithm. SURF

algorithm maintained invariance under scale and affine transformation, and the computing speed was improved by 3 ~ 5 times compared with SIFT, with reduced accuracy.

To solve the above problems, this paper proposed a fusion image registration algorithm based on SFIT and image pyramid, and the process is shown in figure 1:
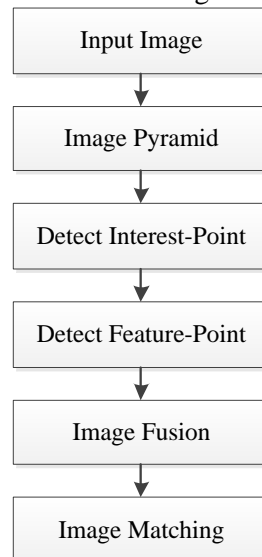
Input Image
↓
Image Pyramid
↓
Detect Interest-Point
↓
Detect Feature-Point
↓
Image Fusion
↓
Image Matching

Figure 1. The process of image registration

## 2. IMAGE PYRAMID

The Image pyramid is an image sequence in which each image consists of a low-pass filter and a secondary sampling sample of its precursor. The Gaussian pyramid and Laplacian pyramid are the main image pyramids in this paper. The Gaussian pyramid is a sequence of images in which each level of the Gaussian pyramid is a copy of the previous level of low-pass filtering [8]. Laplacian pyramid is used to reconstruct the upper layer of unused image from the lower layer image of pyramid. In digital image processing, Laplacian pyramid is also called predictive residual, which can restore the image to the maximum extent.

*2.1 Gaussian pyramid decomposition*

Gaussian pyramid is a series of sub-images obtained through continuous Gaussian low-pass filtering and down-sampling of alternate rows and columns. The bottom layer is the original image, and the size of the remaining images in each layer is a quarter of that in the next layer. In other words, the Gaussian pyramid of the K layer can obtain the Gaussian image of K+1 layer through smoothing and sub-sampling. Let the source image be G0 (m≤ M, n≤ N), M and N are the number of rows and columns of the image. The original image is the bottom layer of the Gaussian pyramid, and the Gaussian pyramid on the L layer is generated by equation (1). According to this formula, G1、G2、G3…GN can be calculated successively to constitute the Gaussian pyramid.

$$G_I(i,j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} \varpi(m,n) G_I(2i+m, 2j+n) \ (1 \leq I \leq N, 0 \leq i \leq R_L, 0 \leq j \leq C_L) \quad (1)$$

Where N is the layer number of the top layer of Gaussian pyramid image; RL and C L are respectively the number of rows and columns in the L layer of Gaussian pyramid image: $\varpi(m,n)$ is a template, which is actually a Gaussian low-pass filter.

*2.2 Laplacian pyramid decomposition*

The Laplacian pyramid is the difference image between the Gaussian pyramid and its upper layer expanded by interpolation, which reflects the information difference between the two levels of the

Gaussian pyramid and is actually the detailed part of the image. The L-layer sub-image GL of the Gaussian pyramid is interpolated to obtain the enlarged image GL*, so that the size of GL* is the same as that of GL-1, and can be expressed as:

$$G^*_l(i, j) = 4 \sum_{m=-2}^{2} \sum_{n=-2}^{2} \varpi(m, n) G_1(\frac{i+m}{2}, \frac{j+n}{2}), (1 \leq l \leq N, 0 \leq i \leq R_L, 0 \leq j \leq C_L) \quad (2)$$

Where

$$G^*_1(\frac{i+m}{2}, \frac{j+n}{2}) = \begin{cases} G_1(\frac{i+m}{2}, \frac{j+n}{2}) & \text{When } \frac{i+m}{2}, \frac{j+n}{2} \text{ are integers} \\ 0, \text{others} \end{cases}$$

Let

$$LP_l = G_l - G^*_{l+1}, 0 \leq l \prec N$$
$$LP_N = G_N, l = N \quad (3)$$

In formula (3), N is the layer number of the top layer of the Laplace pyramid, and $LP_l$ is the image of the L layer decomposed by the Laplace pyramid. LP0, $LP_1$, …, $LP_l$, …, $LP_N$ can be obtained from formula (2) and (3), and the pyramid formed by the sub image series is the pyramid of Laplace. Each layer of the image is the difference between the original layer of the Gaussian pyramid image and the upper layer image after interpolation, which is actually equivalent to band pass filtering.

From formula (3):

$$G_N = LP_N, \text{ when } l = N \quad (4)$$
$$G_l = LP_l + G^*_{l+1}, \text{ when } 0 \leq l \prec N \quad (5)$$

Formula (4) shows that the Laplace pyramid can be recurred from the top layer to the top layer according to formula (4) and formula (5), and the corresponding Gaussian pyramid can be obtained, and the original image G0 can be finally obtained.

## 3. SIFT SCALE INVARIANT FEATURE TRANSFORM METHOD

The SIFT feature is a local feature of the image, which maintains the invariance of scale, rotation, brightness, etc., as well as the stability of Angle change, affine transformation and noise, but the time performance is low and the matching accuracy is not enough. The SIFT image registration basic steps are as follows: 1) feature point extraction; 2) generate feature description operator;3) feature point matching.

### 3.1 Extracting feature points

Generate the DOG scale space. The scale space can be obtained by convolution of Gaussian function and image. The scale space of image is defined as:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (6)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (7)$$

Where $I(x, y)$ is the original image; $\sigma$ is variable kernel，it is a scale factor; * is the convolution operation in x and y, G is a Gaussian smooth kernel function.

The height check scale space $D(x, y, \sigma)$ is obtained by subtracting the images of the upper and lower layers adjacent to the same scale in the Gaussian image pyramid. The formula is:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (8)$$

Where the k is a constant factor and represents the scale of two adjacent scale Spaces

*3.2 Generating feature descriptors*

The region with a size of 16×16 pixels is taken as the center of the feature point, and the region is divided into 4×4 sub-blocks. The histogram of 8 directions of each sub-block can be counted to obtain a seed point. Each feature point is composed of 4×4 seed points, and each seed point has 8 directions, so it forms a feature vector of 4×4×8=128 dimensions, which has rotation invariance, scale invariance and so on. In order to ensure that the vector has certain illumination invariance, it is necessary to normalize it.

*3.3 Key Point Matching*

The SIFT matching algorithm judges the similarity of feature descriptors according to the Euclidean distance between them. The smaller the Euclidean distance is, the higher the similarity will be; otherwise, the lower the similarity will be. In addition, in order to reduce mismatching and improve the accuracy of matching, the ratio refinement method is also adopted to purify the matching results. After the ratio purification, the Random Sample Consensus (RANSAC) algorithm and the Contrario RANSAC ( AC-RANSAC) method are generally used to exclude mismatched points to further improve the matching accuracy.

## 4. PROPOSED ALGORITHM

The algorithm process proposed in this paper is as follows:

   1). The original high-resolution input image was filtered with Gaussian at different scales to obtain the Gaussian pyramid image;

   2). The SIFT algorithm was used to extract feature points of sub-sampled images.

   3). Filter the feature points of sub-sampled images to eliminate the key points with low contrast, and remove the unstable edge response points generated by the algorithm by setting the threshold;

   4). Get the feature points of the original high-resolution image through the feature points of sub-sampling, and then conduct image registration.

## 5. RUSULTS AND DISSCUSSION

The computer configuration of this experiment is: Intel core i5-7600 processor, 3.50GHz frequency, 32GB memory. Programming and running platform is MATLAB R2014a. The image adopted in this paper is from reference 3, and the gaussian pyramid of the original input image is established as shown in the figure below [9]:
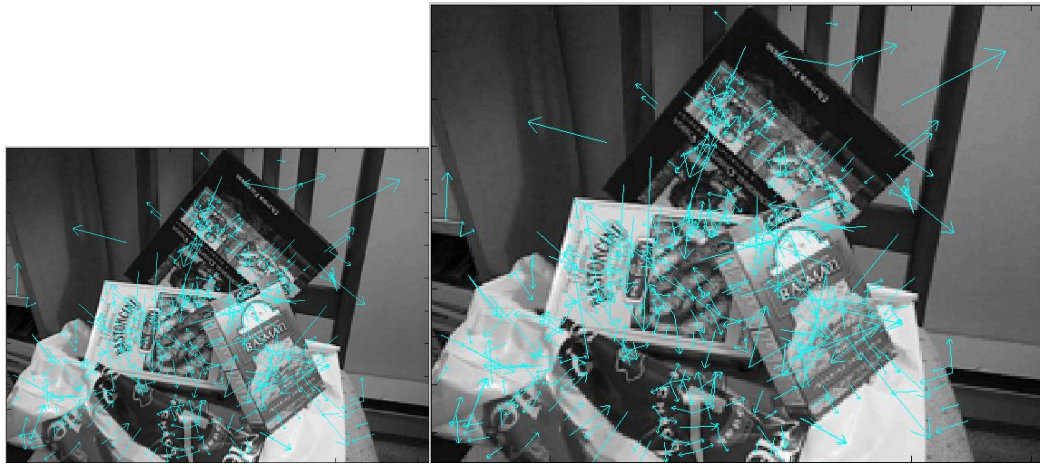


Figure 2. Image pyramid

Figure 3. Image keys used for matching of different algorithm.

Table 1. Comparison table between traditional SIFT algorithm and the algorithm in this paper

|                         | Time/s | Number of feature points |
|-------------------------|--------|--------------------------|
| SIFT Algorithm          | 0.584  | 1021                     |
| Improved SIFT Algorithm | 0.388  | 692                      |

It can be seen from table 1 that the speed of feature point extraction of this algorithm is 150% of that of traditional SIFT algorithm, and the number of feature points extracted is 67.8% of that of traditional SIFT algorithm.

## 6. CONCLUSION

In this paper, an improved algorithm based on SIFT algorithm and image pyramid is proposed. With the improvement of image resolution, the traditional SIFT algorithm needs more time to find the feature points of the image, but it also takes very little time. In this paper, through comparison and analysis of the algorithm in this paper and the traditional SIFT algorithm, the experimental study found that the speed of the algorithm in this paper to find feature points is 150% of the traditional SIFT algorithm.

**References**

[1]   Meng W, Hao L, TAO D, et al. Multimodal Graph-Based Reranking for Web Image Search[J].
        IEEE Transact.
[2]   ZITOVÁB, FLUSSER J. Image registration methods: a survey[J]. Image and Vision Computing,
        2003, 21 (11).
[3]   Brown M, Lowe D G. Automatic Panoramic Image Stitching using Invariant Features[J].
        International Journal of Computer Vision, 2007, 74(1):59-73.
[4]   Harris C，Stephens M . A combined corner and edge detector[C]／／ Alvey Vision Conference，
        IEEE，1988：147--151 .
[5]   LOWE D G. Distinctive Image features from scale-invariant keypoints [J]. International Journal
        of Computer Vision, 2004, 60 (2) :91-110.
[6]   HEN C C, HSIEH S L. Using binarization and hashing for efficient SIFT matching[J].Journal of
        Visual Communication and Image Representation, 2015, 30 (C) :86-93.
[7]   Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, "SURF: Speeded Up Robust
        Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359,
        2008.
[8]   Burt P J, Adelson E H. The Laplacian Pyramid as a Compact Image Code[J]. Readings in
        Computer Vision, 1987, 31(4):671-679.

[9]  Lowe D G. Object Recognition from Local Scale-Invariant Features[C]// iccv. IEEE Computer Society, 1999.