# Image Stitching Algorithm Based on SURF and Wavelet Transform

Jinxin Ruan[1], Liying Xie[2*],Yuyan Ruan[3], Lindong Liu[1], Qiang Chen [1], Qian Zhang[1]
1.Department of Computer Science; 2.Department of physics and Information Engineering 3. College of Science and Engineering
Guangdong University of Education[1,2],City University of Hong Kong[3]
Guangzhou, China[1,2], Hong Kong, China[3]
ruanjinxin@163.com

*Abstract*—**With the development of image stitching and its wide application, image stitching has become an important and heated topic in image processing. An effective image stitching algorithm based on SURF feature matching and wavelet transform image fusion is proposed in this paper. Firstly, SURF feature points in the two adjacent images are extracted and matched. Then rapid and accurate image registration can be achieved by adopting the improved RNASCA algorithm to removing the mismatched feature point pairs. The multi-resolution decomposition of the overlapping region is processed by Wavelet Transform. Then the multi-scale image fusion is processed by fade-in and fade-out in order to eliminate the stitching seam better. Experiments show that the fusion results of the overlapping region are natural and there also is a certain robustness for translation, rotation, scale and luminance variant.**

***Keywords-Image stitching; SURF; RNASCA; Image Fusion; Wavelet Transform***

## I. INTRODUCTION

Image stitching refers to stitching two or even more images which have overlapping regions into an image with larger viewing angle and higher resolution. And image stitching technology can effectively solve the problem of small viewing angle of imaging equipment, satisfying the demand for large view images. Therefore, image stitching has wide application prospect and practical value in military fields like military reconnaissance, etc. and civil fields such as safety monitoring, intelligent traffic, emote sensing image processing, satellite image processing, medical image processing, virtual reality and so on. Now, image stitching has become one of the most heated and important topics in image processing[1].

The process of image stitching includes image acquisition, image preprocessing, image registration, image fusion and image output. Only when the overlapping regions of two adjacent images are registered accurately can the stitching quality be guaranteed. Therefore, image registration is the crucial step in image stitching. Another important step is image fusion, and it is to eliminate the discontinuity caused by the difference of image color and brightness in the overlapping region so as to realize natural fusion and get seamless results.

The corresponding author: Liying Xie (332574459@qq.com)

At present, common image registration algorithms can be roughly divided into three categories: based on gray information[2], based on transform domain[3] and based on feature[1]. The image registration algorithm based on gray information directly uses the gray information to search the optimal value of the two images similarity and determine the overlapping regions of the two images. And this algorithm is simple, but the computation is large and it is easy to be affected by the gray change . The image registration algorithm based on transform domain is to transform the image from spatial domain into frequency domain by Fourier transform. So it is robust to image registration for translation, rotation and scaling. In the image registration algorithm based on feature, the representative feature information of two images, such as corners, edges and contours, is extracted firstly. After matching the image features, the registration mapping relation between the two images is established finally. The registration results of this algorithm are affected by feature extraction and mismatching. But it has the advantages of high calculation speed and good robustness of rotation, scaling and even affine. Therefore, it has become the mainstream algorithm so far.

SIFT(Scale-Invariant Feature Transform, SIFT)[4] and SURF(Speeded Up Robust Features, SURF)[5] are two of the well-known registration algorithms based on features. The SIFT feature, which has relatively strong robustness to rotation, scale and illumination, has aroused the interest and research concern of many researchers, and many improved algorithms have been proposed such as PCA-SIFT[6] and ASIFT[7]. Developed on the basis of the SIFT algorithm, the SURF algorithm, which has the similar robustness of the the SIFT algorithm, simplifies the Gaussian second order partial derivative by adopting integral image and box filter to reduce the time of feature extraction. Compared with the SIFT algorithm, it has greatly increased the speed of feature extraction[8]. However, after extracting the feature, mismatching of the feature points still arise when matching the two images by BBF. And RANSAC (Random sample consensus, RANSAC)[9] is commonly used for removing the mismatching.

The common image fusion algorithms include spatial domain fusion and transform domain fusion[10]. Average algorithm, which belongs to spatial domain fusion approach, has advantages of simple algorithm and high calculation speed. But there will be obvious stitching seams in image

978-1-5386-9497-8/18/$31.00 ©2018 IEEE
DOI 10.1109/ICDH.2018.00009

fusion regions. Median filtering algorithm is also a common spatial domain algorithm, and the its visual results are unsatisfactory for still exiting obvious stitching seams. Laplacian pyramid[11] and wavelet transform image fusion algorithm[12] belong to transform domain fusion algorithm. These algorithms decompose the overlapping region image into multiple band-pass images. And then the different smoothing functions are used for processing the each band-pass images. As a result, the better fusing performance can be achieved in both spatial and spectral in the overlapping region.

In order to achieve the rapid and accurate image registration and get better image fusion, an effective image stitching algorithm, which is based on SURF feature matching and wavelet transform image fusion, is proposed in this paper.

## II. Image Registration Based on SURF Feature

Image registration, which is also known as image matching, refers to aligning two or more images according to image space position of the same feature. And the detailed process of image registration based on SURF feature includes SURF feature point detection, identifying dominant orientation , SURF feature vector construction, SURF feature point matching and removing the mismatched feature pairs.

### *A. SURF feature point detection*

Similar to SIFT algorithm, SURF algorithm uses Hessian matrix to detect feature points. And it can be greatly accelerated by using integral image, box filter and Haar wavelet.

Given a point $\mathrm{X}=(x,y)$ in an image $I(\mathrm{X})$, the two-dimensional Gaussian kernel function $g(\mathrm{X},\sigma)$ in $\mathrm{X}$ at scale $\sigma$ is defined as follow.

$$g(\mathrm{X},\sigma)=\frac{1}{2\pi\sigma^2}\exp[\frac{-(x^2+y^2)}{2\sigma^2}] \tag{1}$$

The scale space function of an image $I(\mathrm{X})$ is defined as $L(\mathrm{X},\sigma)$, and $L(\mathrm{X},\sigma)$ is the convolution of a variable-scale Gaussian function $g(\mathrm{X},\sigma)$ with $L(\mathrm{X},\sigma)$ in $\mathrm{X}$.

$$L(\mathrm{X},\sigma)=g(\mathrm{X},\sigma)*I(\mathrm{X}) \tag{2}$$

where $*$ is the convolution operation.

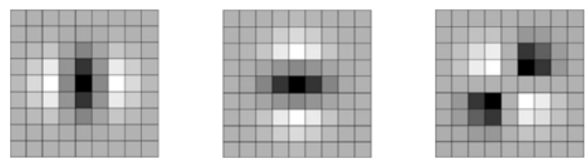
$L_{xx}(\mathrm{X},\sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2 g(\mathrm{X},\sigma)}{\partial x^2}$ with the image $I(\mathrm{X})$ in $\mathrm{X}$, and the $L_{xx}(\mathrm{X},\sigma)$ can be defined as follow.

$$L_{xx}(\mathrm{X},\sigma)=\frac{\partial^2 g(\mathrm{X},\sigma)}{\partial x^2}*I(\mathrm{X}) \tag{3}$$

Similar to $L_{xx}(\mathrm{X},\sigma)$, $L_{xy}(\mathrm{X},\sigma)$ and $L_{yy}(\mathrm{X},\sigma)$ are the convolution of the Gaussian second order derivative with the image $I(\mathrm{X})$ in $\mathrm{X}$. Therefore, given a point $\mathrm{X}$ in an image $I(\mathrm{X})$, the Hessian matrix $H(\mathrm{X},\sigma)$ in $\mathrm{X}$ at scale $\sigma$ is defined as follow.

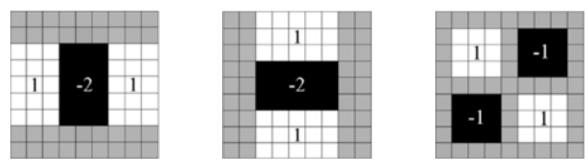
$$H(X,\sigma)=\begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \tag{4}$$

The Gaussian second order partial derivatives are shown in Figure 1. Correspondingly, SURF algorithm adopts box filters to approximate the Gaussian second order partial derivative, and the results are shown in Figure 2. In the direction of $x$ and $y$, the weight of the black regions is -2, and the weight of the white regions is 1. The weight of the grey regions is equal to zero. In the direction of $xy$, the weight of the black regions is -1, and the weight of the white regions is 1, and the weight of the grey regions is zero.



(a)x-direction (Lxx) (b) y-direction (Lyy) (c)xy-direction (Lxy)

Figure 1.The Gaussian second order partial derivative




(a)x-direction (Dxx) (b)y-direction (Dyy) (c)xy-direction (Dxy)

Figure 2. The approximation for the second order Gaussian partial derivative by using box filters

Keeping the original image size, SURF algorithm only resize the box filters, and $D_{xx}$, $D_{xy}$, $D_{yy}$ can be got at different scales space by convoluting the fliters of diferet sizes and the original image. $H_{approx}$ is the approximate matrix of Hessian matrix. Then $H_{approx}$ can be defined as follow.

$$\det(H_{approx})=D_{xx}D_{yy}-(\omega D_{xy})^2 \tag{5}$$

Where $\omega$ is the weigh of the filter responses, which is used for balance the error between exact value and approximate value. In general, $\omega$ equals 0.9.

The approximation of Hessian determinant of every pixel in image $I(\mathrm{X})$ at different scale spaces can be calculated by using the expression of (5). The local extremum, which is considered as the candidate for feature point, can be found by applying non-maximum suppression in the different scale spaces. Then its location and scale information will be recorded and the scale-invariant feature detection will be realized.

### B. Identifying dominant orientation

To estimate the dominant orientation of the feature point, SIFT Algorithm is identified by calculating gradient histogram in feature point neighborhood, however SURF Algorithm is identified by calculating the Haar wavelet responses within a circular neighbourhood of feature point. In detail, at a certain scale, the Haar wavelet response in x or y direction are calculated within a 60-degree, whose center is around the feature point and radius is 6s. And a local orientation is defined as the direction of the sector by the sum of x and y direction. Then the responses are summed when sliding the sector. Finally, the direction of the maximum vector in the sector is identified as the dominant orientation of the feature point.

### C. SURF feature vector construction

The process of SURF feature vector construction is shown in Figure 3. A square region, which is centered around the feature point and oriented along the dominant orientation of the feature point, is constructed. And its size is 20s. And it is split into $4*4$ sub-regions. The response $d_x$ presents the Haar wavelet response in horizontal direction, and the response $d_y$ presents the Haar wavelet response in vertical direction. Then $\sum d_x$ , $\sum |d_x|$ , $\sum d_y$ , and $\sum |d_y|$ are calculated in every region and a four-dimensional vector $V_{sub}$ is formed.

$$V_{sub} = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right) \tag{6}$$

When all the vectors in $4*4$ sub-regions of each feature point are concatenated, a 64-dimensional vector, which is the SURF feature vector of this feature point, is to be constructed.

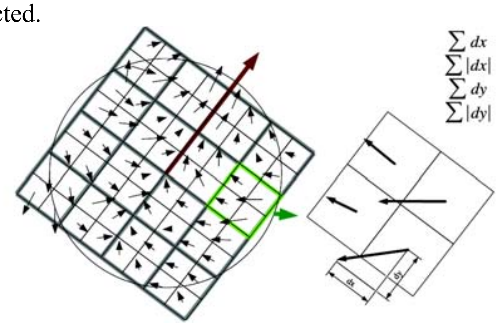



Figure 3. The process of constructing SURF feature vector [5]

### D. SURF feature point matching and removing the mismatched feature pairs

Feature point matching is to find the corresponding feature points in the images, which are to be registered. And the feature point pairs is constructed. The transformation relation between the point $(x_i, y_i)$ in one image and the point $(x_i^{'}, y_i^{'})$ in the other image can be estimated by the correspondence of the feature point pairs. The transformation relation, which is also called image transformation model, can be defined as follow.

$$\begin{bmatrix} x_i^{'} \\ y_i^{'} \\ 1 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \tag{7}$$

In this paper, BBF (Best Bin First, BBF)[4] algorithm， which is improved based on $K-d$ tree algorithm and is effective algorithm, is adopted for matching the the feature points. In BBF algorithm, matching feature points is achieved by searching the nearest neighbourhood feature point of each feature point. However, a threshold value, which is required to be set to select feature points during the searching process, will cause some mismatching of feature points. Therefore, data fitting algorithm is needed to remove those mismatched feature point pairs. RANSAC algorithm is an relatively effective data fitting algorithm, which adopts random sampling to removing the error data. Additionally, in RANSAC algorithm, optimal parameters of the corresponding model can be obtained when clustering is completed. In this paper, improved RANSAC algorithm[13] is adopted to remove the error SURF pairs, meanwhile, the parameters of the image transformation model can be obtained.

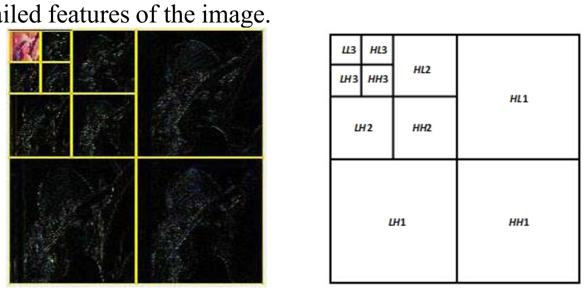
## III. MULTI-RESOLUTION IMAGE FUSION BASED WAVELET TRANSFORM

Image fusion refers to fusing the two registered images into one image. For Wavelet transform has the characteristics of time-frequency localization, it can extract effective information from images. Therefore, it has better fusion performance compared with the Laplacian pyramid algorithm. In practical application, the another algorithm of fade-in and fade-out image fusion is relatively easy and achieves better stitching results. Therefore, multi-resolution image fusion based wavelet transform is adopted in this paper to process the overlapping regions, combining with fade-in and fade-out image fusion.

### A. Wavelet Transform

In Wavelet Transformation, the signal can be gradually multi-scale refined by scaling and translation operation step by step. Therefore, Wavelet transform can focus on any detail of the signal and it is called "mathematical microscope".

In image processing applications, the image can be decomposed into low frequency image and high frequency image after wavelet decomposition, and Low frequency images can continue to be decomposed step by step. And the decomposed sub-images at all levels contain the spatial structure information of the original image. As shown in figure 13, it is the multi-resolution image obtained by three-level 2-D discrete wavelet transform (discrete Wavelet Transform， DWT). LL is the low frequency sub-image, which keeps most information of the image. LH, HL, and HH are high frequency sub-images of vertical, horizontal, and diagonal directions respectively, which denote the detailed features of the image.




(a)The multi-resolution image of lena (b) The image decomposition

Figure 4. The multi-resolution image after three-level 2-D DWT

Generally, some edges or details in the images exist within a certain scale. In other words, the images at any scale can ’ t clearly reflect all the features and detailed information. The wavelet decomposition is multi-resolution and multi-scale. Hence, the multi-scale decomposing process can be regarded as the process of the multi-scale edge information extraction. Meanwhile, the multi-scale wavelet decomposition has directionality. When applying wavelet transform to image stitching, edges and details of various sizes and directions can be stitched at different scales so that better stitching results can be obtained.

### *B. Fade-in and Fade-out Image Fusion*

Given a point $X=(x,y)$ is a point in image. $I_1(X)$ and $I_2(X)$ are the two images to be registered, and $I(X)$ is the fused image. In fade-in and fade-out image fusion algorithm, $I(X)$ can be defined as follow.

$$I(x,y)=\begin{cases} I_1(x,y) & (x,y)\in I_1 \\ \beta I_1(x,y)+(1-\beta)I_2(x,y) & (x,y)\in I_1\cap I_2 \\ I_2(x,y) & (x,y)\in I_2 \end{cases} \tag{8}$$

Where $\beta=\frac{x_{\max}-x}{x_{\max}-x_{\min}}$ , $x_{\max}$ and $x_{\min}$ are the maximum and minimum respectively in $x$ direction of the overlapping region, and the width of the overlapping region is $T=x_{\max}-x_{\min}$ .

When $\beta$ changes from 1 to 0, the image of overlapping region transits from $I_1$ to $I_2$ accordingly. Hence, the fade-in and fade-out smooth transition of images can be obtained and the stitching seam can be better eliminated.

### *C. Multi-resolution Image Fusion Algorithm based on Wavelet Transform*

Multi-resolution image fusion algorithm based on Wavelet Transform is adopted to image fusion of the overlapping region. And the detailed algorithm flow is shown in Figure 9. Firstly, the overlapping region of the two adjacent images to be stitched is extracted . Then sub-images of multi-resolution are obtained by multi-level 2-D wavelet decomposition. And the different regions of the sub-images are fused by fade-in and fade-out algorithm with different the width T. Finally, the image is reconstructed by 2-D inverse wavelet transform and image fusion is achieved.

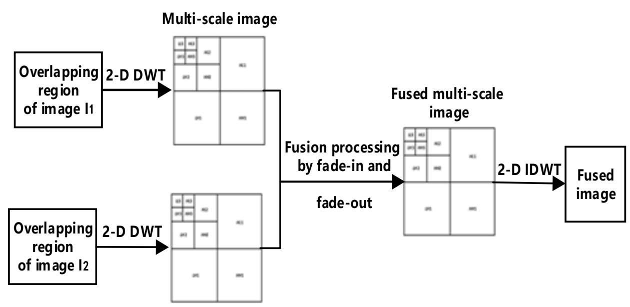



Figure 5. Detailed algorithm flow of overlapping region fusion based on DWT

## IV. EXPERIMENTAL RESULTS AND ANALYSIS

The images used in this paper were taken by Nikon COOLPIX P900s digital camera. The image format is RGB color and the image size is $640*480$ . The following five pictures, which are shown as Figure 6, were taken by the camera. During photographing, the camera is fixed with the tripod, however, the pictures were taken with some difference of the location, the lights, the angle and the focus. As a result, the chosen pictures varied in light, visual angle and scale.

The experiments were carried on the computer which has Intel Core(TM) i7-4712MQ CPU@2.30GHz, 12GB memory and NVIDIA GeForce GT 720M；, programming based on Microsoft Visual Studio 2015 and OpenCV 3.4.0. And the computer systems is Microsoft Windows 10 Pro 64 Bit System.

From the result of image stitching in Figure 7, the profile height of each part of the picture varies because of the differences in the focal length, resolution and angle of view between different pictures. The overlapping regions of the stitching image are uniform in the luminance and colour. Because there are almost no ghosting and seams, the stitched image achieves good performance. And it also demonstrate that the algorithm proposed in this paper is robust to the translation, scale, rotation and luminance change of stitching images.

(a) (b) (c) (d) (e)

Figure 6. The five images to be stitched

Figure 7. The image stitching result

## V. CONCLUSIONS

This paper analyzes the key technology of image stitching which includes image registration technology and image fusion technology. And the strong emphasis is laid on the study of the image registration algorithm based on SURF feature and multi-resolution image fusion algorithm based Wavelet Transform. Experiment results show that the image stitching algorithm proposed in this paper is effective and the image stitching results are satisfying.

## ACKNOWLEDGMENT

The research was granted by the National Natural Science Foundation of China(No.61772140), Science and Technology Planning Project of Guangdong Province(No.2017A010101021), key disciplines of network engineering of Guangdong University of Education, Appropriative Researching Fund for Professors and Doctors, Guangdong University of Education (No.2015ARF25), Researching Fund for Professors and Doctors, Guangdong University of Education (No.2015ARF13).

## REFERENCES

[1] Adel E, Elmogy M, Elbakry H. Image Stitching based on Feature Extraction Techniques: A Survey[J]. International Journal of Computer Applications, 2014, 99(6):1-8.

[2] Liu J, He-Lei W U. A Tomography Image Re-Registration Algorithm Based on Frame and Gray Information[J]. Computer Engineering & Science, 2012, 34(12):110-114.

[3] Wang Q, Wang Z. Color image registration based on quaternion Fourier transformation[J]. Optical Engineering, 2012, 51(5):7002, doi: 10.1117/1.OE.51.5.057002.

[4] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[5] Bay H, Tuytelaars T, Gool L V. SURF: Speeded Up Robust Features[C]// European Conference on Computer Vision. Springer-Verlag, 2006:404-417.

[6] Y. Ke and R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors," Computer Vision and Pattern Recognition, pp. 511-517, 2004.

[7] Morel J M, Yu G. ASIFT: A New Framework for Fully Affine Invariant Image Comparison[J]. Siam Journal on Imaging Sciences, 2010, 2(2):438-469.

[8] Luo J, Gwun O. A Comparison of SIFT, PCA-SIFT and SURF[J]. International Journal of Image Processing, 2013, 3(4):143-152.

[9] Hast A, Nysjö J, Marchetti A. Optimal RANSAC—Towards a Repeatable Algorithm for Finding the Optimal Set[J]. Journal of Wscg, 2015, 21(1):21-30.

[10] Liu C, Jin L, Tao H, et al. Multi-Focus Image Fusion Based on Spatial Frequency in Discrete Cosine Transform Domain[J]. IEEE Signal Processing Letters, 2014, 22(2):220-224.

[11] Wang W, Chang F. A Multi-focus Image Fusion Method Based on Laplacian Pyramid[J]. Journal of Computers, 2011, 6(12):2559-2566.

[12] Jadhav S. Image Fusion Based On Wavelet Transform[J]. International Journal of Engineering Research, 2014, 3(7):442-445.

[13] Zhang Z, Kuang L, Han X, et al. An Improved RANSAC Algorithm for Image Mosaic[J]. Computer Measurement & Control, 2014.