Research on Fundus Image Registration And Fusion Method Based On NSCT And Adaptive PCNN

Jun Wu¹,², Shuya Song¹,²*, Dongxia Zhang¹,², Song Yang¹,²
¹ Tianjin Polytechnic University, School of Electronics and Information Engineering, Tianjin 300387, China
² Tianjin Key Laboratory of Optoelectronic Detection Technology and System, Tianjin 300387, China
Corresponding author: Jun Wu

Abstract: The fusion of the fundus image can show lesions and tissues in multi-modalities of images comprehensively, and is of great value of application in the diagnosis of fundus diseases for doctors. In this paper, we propose a method of fundus image registration and fusion with the combination of NSCT (Nonsubsampled Contourlet) and adaptive PCNN (Pulse Coupled Neural Network). The specific process is: firstly, two images are registered to eliminate the spatial differences between the source images to extract SURF (Speeded Up Robust Features) as the feature point, and then the feature vector is calculated by feature points description, the nearest neighbor and the next nearest neighbor distance ratio method are used to realize the initial matching of the feature points, the RANSAC (Random Sample Consensus) method is used to remove the mismatched point pairs. Finally, the transformation parameters between the images are calculated to complete the registration of source images by the spatial transformation. For integration the of the two images after the registration, the specific process is: The low frequency and high frequency sub-band of the image to be fused is got by NSCT decomposition, the low frequency sub-band is fused by the regional energy. The high frequency sub-band use the simplified PCNN model to study and is fused based on the number of times the image pixels are fired. The experimental results show that this method is better than other representative methods in the fusion result of fundus image, the fusion image synthesizes the image information and clarifies the performance of the details, provides an effective reference for the clinical diagnosis of fundus diseases.

Keywords: Fundus image, NSCT, regional energy, simplified PCNN

Date of Submission: 17-01-2018
Date of acceptance: 31-01-2018

1. INTRODUCTION

Image fusion is to combine two or more images from the same scene into a single image, and the fused image contains more information than a single original image[1]. In recent years, medical image fusion technology has been widely attention, which makes the clinical diagnosis and treatment more comprehensive and accurate. In terms of fundus images, Fundus diseases occur not only affect people's normal life, but also may lead to sharp decline in vision or even blindness. And many systemic diseases such as diabetes, hypertension, hematological diseases and so on will also cause some changes in the characteristics of the fundus images, resulting in abnormal morphology of fundus tissue information, clinical manifestations in the fundus position include Optic disc lesions, micro-aneurysms, hemorrhagic spots, bright yellow exudates, vascular tortuosity, neovascularization, and expansion of capillaries etc. color fundus images and fluoroscopy fundus images can provide different angles of reference information, in the diagnosis of each have their own advantages and disadvantages. Fluorescence fundus images showed a high fluorescence effect, can clearly observe the vascular lesions and obstruction micro-aneurysm cases, but bleeding spots and exudate lesions image is not conducive to human eye observation in the fluorescence. Color fundus images can compensate for this limitation, it is able to a more comprehensive reflection of the fundus information, observe the bleeding spot and exudation lesions clearly, but microvascular tumor and vascular lesions contrast is low in the color fundus image, the characteristics are not obvious. So, the color fundus image is complementary to the fluoroscopic fundus image. moreover, the clinical manifestations of hypertensive patients in the fundus of the retina is generally narrowed arteries, the normal arteries and veins ratio of 2:3, then the ratio can be changed to 1:3 or even 1:4, can also produce local stenosis. When the blood pressure continues to rise, will form atherosclerosis, almost no blood vessels, but the fluorescence imaging can fill the blood vessels, so that the performance of clear; some blood diseases such as anemia can lead to retinal lesions, the clinical manifestations of bleeding, seepage Evidence of pale discoloration of the optic disc and changes of the blood vessels, the disc images appear clear in the color fundus images, corresponding to the fluorescence images of the fundus of the eyes was black fluorescence shading, so the fusion of color fundus
images and fluorescence images of the fundus, the effective formation of information. Complementary, more comprehensive and intuitive reflect the characteristics of lesions and organizations for the prevention and treatment of some diseases has important application value.

In practical applications, image registration and image fusion technology are inseparable. In general, the imaging sensor to obtain the source image will be different in position and attitude, which makes the source image inevitably exist differences geometry space of translation, rotation and scale expansion etc. This difference has great influence on the effect of image fusion. Therefore, before the image fusion of the source image, it is necessary to do the geometric alignment by image registration for eliminate the spatial differences between source images which comes from different angles and sensors of the same scene. Image registration methods can be divided into region based methods[2] and feature based methods[3-4]. Region based method deals with objects as pixels, and uses image templates to process registration. The differences between different methods are reflected in the design of templates and registration rules; Feature-based method by extracting spot, lines, contours and other significant features in the source image as a unit for feature matching.

The image fusion method can be divided into image fusion based on spatial domain and image fusion based on transform domain. Image fusion based on spatial domain is the fusion of images directly in the gray space of image pixels, such as weighted average fusion method, IHS spatial fusion method and principal component analysis fusion method etc.. The method of spatial domain fusion is simple and the method has low complexity, but the method is difficult to achieve satisfactory fusion effect in most applications. At present, the fusion method based on transform domain is the focus of the research. Discrete wavelet transform (DWT) as an image multi-scale geometric analysis tool, which has good time-frequency local analysis, but the orthogonal wavelet transform based on the Mallat method can only decompose the finite direction information, and does not have translation invariance, resulting in ringing effect in the reconstructed image[5]. At present, scholars have proposed many multi-scale geometric analysis methods based on the wavelet transform theory; Candes and Donoho proposed the Ridgelet Transform[6], Curvelet transform[7], Do and Vetterli proposed Contourlets transform (CT)[8], Contourlets transform is a "true" two-dimensional image representation. Compared with the traditional DWT transform, Contourlets transform not only has the characteristics of multiresolution and time-frequency local analysis, but also has multi-directional and anisotropy, which can effectively capture the smooth contours in the image, and effectively describe the directional texture of the image Information. However, due to the up and down sampling operations in decomposition and reconstruction, Contourlets transform does not have the translation invariance, prone to spectral aliasing, resulting in the emergence of pseudo-Gibbs phenomenon. In order to compensate the pseudo-Gibbs phenomenon, Cunha proposed Nonsubsampled contourlet (NSCT)[9], This method inherits the Contourlet transform, avoids the lower sampling operation, has translational invariance, can better express the contour information and edge information of the image, and overcomes the phenomenon of pseudo Gibbs. So, NSCT for image fusion can achieve better fusion effect. After multi-scale decomposition, the corresponding fusion rules should be developed for different types of images. In view of the multi focus images, Song Ruixia[10] proposed to use the weighted average of spatial frequency, variance, and improved Laplasse energy to fuse the low-frequency sub-bands in the NSCT domain, using local texture features to fuse high frequency sub-bands. For brain medical images, Dai Wenzhan proposed to use regional energy and mean gradient to fuse the low frequency sub-bands in the NSCT domain, the high frequency sub-band is fused by Region Laplasse energy, directional contrast, and PCNN. Yang Licai[12] proposed the medical image is decomposed by wavelet packet, and the wavelet coefficients and the decomposed sub-images are processed by adaptive operator. The medical fusion image is obtained by wavelet packet reconstruction., Li Xine[13] proposed to use regional energy to fuse the low frequency sub-bands in the NSCT domain, using the band-pass direction sub-band coefficient as the external input excitation of the PCNN to obtain the ignition map to fuse the high frequency sub-band.

Based on the above study, this paper considers the traditional image feature having limitations in color fundus images and fundus image fusion with complex background, proposed a fusion method of fundus image fusion based on NSCT and adaptive PCNN. Firstly, eliminating the spatial difference of color fundus images and fluorescein fundus images that comes from the same person with the same eye shot at the same time. Then, the two fundus images were decomposed by NSCT to obtain the low frequency sub-band and high frequency sub-band of two images. According to the characteristics of the low frequency sub-band to determine the image profile, using method of regional energy to fuse the low frequency sub-bands of two images. According to the characteristics of high frequency sub-bands reflect the details of the image, using the improved simplified PCNN network and aiming at the difficulty of the problem of determining the parameter of PCNN network, an adaptive image fusion method based on improve the Laplacian energy is proposed. It can adaptively set the parameter of PCNN and fuse the high frequency sub-bands of two images according to the number of fires. the three channels are fused respectively, the fusion image synthesizethe image information and clarify the performance of the details.

www.ijeijournal.com
II. REGISTRATION

In this paper, an image registration method based on accelerated robust feature (SURF) [14] feature points is proposed. The fluorescence image is selected as the reference image, and the color fundus image is the floating image to be registered. Firstly, extracting feature points in different scale spaces. Then, the feature vectors are computed by feature description, and the initial matching of feature points is achieved by using the ratio of the nearest neighbor to the next nearest neighbor. Then, using the RANSIC method to remove false match pairs. Finally, the transformation parameters between the images are calculated, and the spatial transformation of the color images is completed. The experiments show that this method can realize automatic registration of images under different pathological conditions. 2.1 Feature extraction

The detection of SURF feature points is based on Hessian matrix and scale space theory, using the Hessian matrix to detect the extreme points in the image scale space as candidate feature points. In image $f$, given any $p = (x, y)$, in position $p$ and scale $\sigma$, the second order Hessian matrix $H(p, \sigma)$ can be defined as:

$$H(p, \sigma) = \begin{bmatrix} I_{xx}(p, \sigma) & I_{xy}(p, \sigma) \\ I_{yx}(p, \sigma) & I_{yy}(p, \sigma) \end{bmatrix} \quad (1)$$

Where: $I_{xx}$, $I_{xy}$, $I_{yx}$, and $I_{yy}$ express the convolution of higher-tow-order partial derivative in spot $p$ with image $f$.

In order to detect the feature points in different scale spaces, using Gauss difference function to construct the scale space Pyramid to guarantee method has scale invariance. In order to improve the calculation speed of convolution, using box filter to instead Gauss two derivative by adjusting the size of the block filter. Then, the scale space $D(p, \sigma)$ is constructed by convolution of the original image with block filters of different sizes. Finally, in the three-dimensional space $(x, y, \sigma)$, each feature point is subjected to non maximal suppression operations in $3^3$ neighborhood. By comparing with the surrounding 26 points, the point with the largest response value is selected as the feature point.

III. FEATURE DESCRIPTION

PIIFD(Partial Intensity Invariant Feature Descriptor)is a local feature descriptor proposed by Chen Jian [15] in multi-modal fundus image registration. PIIFD has image rotation invariance, partial intensity, affine transformation, and angle invariance. The PIIFD presentation is based on such image features: (1) The structure area of an image corresponds to a similar contour in the corresponding region of the other image. In this paper, the image contour extraction is simplified to extract the image gradient. (2) The gradient direction of the corresponding position of the two multimodal images points in the same direction or the opposite direction. In order to extract PIIFD, firstly, the gradient size and direction are sampled in the neighborhood of the feature point. In the scale space of the SURF feature point, taking the point as the center of $40 \times 40$ pixel size of the neighborhood. Then, this neighborhood is divided into $4 \times 4$ sub neighborhood, calculating the pixel gradient in each sub neighborhood and accumulating in the corresponding histogram, a direction histogram covering 0 to 360 degrees, dividing the histogram into uniform 16 directions ($0^\circ$, $22.5^\circ$, $45^\circ$, ..., $337.5^\circ$). Then, Standardizing the size of the gradient, thereby reducing the impact of changes in the size of the gradient. In the neighborhood of the feature point, the strongest 20% of the gradient is represented by 1. Sub strong is represented by 0.75, and so on, the weakest 20% is represented by 0. The second step is to reduce the 16 directions in 8 directions by means of summing in the opposite direction, ($0^\circ$, $22.5^\circ$,$45^\circ$, ..., $157.5^\circ$). All histograms can be represented as:

$$P = \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{bmatrix} \quad (2)$$

represents a histogram containing 8 directions, the PIIFD feature descriptor is expressed as:

$$D_p = \begin{bmatrix} P_1 + \text{rot}(P_1, \pi) \\ P_2 + \text{rot}(P_2, \pi) \\ \alpha[P_3 + \text{rot}(P_3, \pi)] \\ \alpha[P_4 + \text{rot}(P_4, \pi)] \end{bmatrix} \quad (3)$$
Where $P_i$ is the direction histogram of the first row, $P_2$, $P_3$, $P_4$, and so on. $rot(P_i, \pi)$ represents the direction of the histogram rotated $180^\circ$, where:

$$\alpha = \frac{\max(P_i + rot(P_i, \pi))}{\max(P_i - rot(P_i, \pi))}, i = 3, 4 \quad (4)$$

So, the PIIFD feature descriptor is $4 \times 4 \times 8 = 128$ dimensional feature vector, normalized to a unit length, and has the features of no rotation distortion, partial intensity invariance and view invariance.

**Feature Point Matching:** The feature matching is to search for the most similar eigenvectors in the vector space. After the feature description vectors are generated, the nearest neighbor and sub nearest neighbor ratio method is used to realize the initial matching of feature points. RANSIC method is used to realize the precise matching of feature points, when the correct matching points are obtained, the least squares method is used to estimate the parameters of the transform model. Then, the color fundus images are processed by geometric transformation through using the calculated parameters. The two source images and the color fundus images after registration are shown in Figure 2.

![Figure 2](image)

**Figure (A) Color Fundus Image (B) Fluorescein Fundus Image (C) Color Fundus Image After Registration Image Fusion**

The low frequency sub-band and high frequency sub-band of the image are obtained by NSCT transformation of the two images after registration. For the low frequency sub-band, using the method of regional energy and Phase congruency to fuse image. For the high frequency sub-band, the two images are input to the PCNN network optimized by particle swarm to obtain the number of ignition times and according to the number of ignition for fusion. Finally, NSCT inverse transform is used to get the fused image. 3.1 Image decomposition. The NSCT transform inherits the Contourlet transform and avoids the downsampling operation, having translation invariance, which can better represent the contour information and edge information of the image. Contourlets transform is a "true" two-dimensional image representation. NSCT consists of Nonsampled Pyramid (NSP) and Nonsampled Directional Filter Bank (NSDFB). Firstly, NSP is used to perform multi-scale decomposition of the image to obtain a low frequency sub-band and a high frequency sub-band, which of size is the same as the original image. Then, the low-frequency sub-band repeats NSP operation, NSDFB is used to decompose the high frequency sub-band of each stage NSP. We get different scales, different directions of the sub-band coefficient. The block diagram shown in Figure 2.
Low frequency sub-band coefficient fusion: After the image is transformed by NSCT, the low frequency part concentrates most of the energy of the image. The purpose of low frequency fusion is to effectively preserve the information of color fundus images, and to fuse the feature information of fluorescein angiography fundus images. The regional energy can reflect the distribution of the luminance information, and the area energy is the weight of the fusion coefficient to effectively preserve the brightness information of the source image. The energy of the low frequency sub-band centered on \((i, j)\)is calculated, define as follows:

\[
E_{LM}^{M,N}(i, j) = \sum_{m=0}^{M} \sum_{n=0}^{N} (C_{LM}^{M,N}(i + m, j + n))^2
\]  

(5)

\(C_{LM}^{M,N}\) Represents the low frequency coefficient of point \((i, j)\), the area size is \(S \times T\).

When the region energy of image M are greater than the image N, it is shown that the information of the image M is significant at this time, and the coefficient of image after fusion selects low-frequency coefficient of image M. Conversely, Selecting the low frequency coefficients of image N as the coefficient of image after fusion.

In addition to the above two cases, regional energy is used to weighted fusion can preserve most of the information of color fundus images.

Define the region energy duty ratio of image M and N:

\[
K_M = \frac{E_M^L (i, j)}{E_M^L (i, j) + E_N^L (i, j)}
\]

(6)

\[
K_N = \frac{E_N^L (i, j)}{E_M^L (i, j) + E_N^L (i, j)}
\]

Finally, the low-frequency sub-band coefficients of the fused image are determined:

\[
C_L^F (i, j) = \begin{cases} 
C_M^L (i, j), & E_M^L (i, j) > E_N^L (i, j) \\
C_N^L (i, j), & E_M^L (i, j) < E_N^L (i, j) \\
K_M C_M^L (i, j) + K_N C_N^L (i, j), & \text{other}
\end{cases}
\]

(7)

Where: \(E_M^L (i, j), E_N^L (i, j)\)respectively represents the regional energy of the low-frequency subband M and N at the point \((i, j)\).

3.3 High frequency sub-band coefficient fusion

3.3.1 PCNN principle

PCNN is widely used in image processing due to its global coupling and dynamic pulse propagation characteristics[18]. A large number of nonlinear modulation mechanisms and multiple leakage integrators exist in the classical PCNN model, which of parameter setting is complex and inconvenient for practical application. So, in this paper, a simplified PCNN model is used, its mathematical equation is described below:
Research On Fundus Image Registration And Fusion Method Based On NSCT And Adaptive PCNN

\[ F_y(n) = S_y(n) \exp(-\alpha_y) + V_y \sum W_{al} Y_a(n-1) \]  
\[ U_y(n) = F_y(n) \times (1 + \beta L_y(n)) \]  
\[ \theta_y(n) = \exp(-\alpha_y) \theta_y(n-1) + V_y Y_y(n-1) \]  
\[ Y_y(n) = \begin{cases} 1, & U_y(n) \geq \theta_y(n) \\ 0, & \text{else} \end{cases} \]

Where: \( S_y \) is an external input stimulus; \((i, j)\) represents the coordinates of the pixel or the coordinates of the neurons; \( F_y \) is the feedback input of neurons; \( L_y \) is the link input of the neuron; \( U_y \) is an internal activity item; \( \theta_y \) is a dynamic threshold; \( Y_y \) represents the pulse output of PCNN; \( n \) is the number of iterations; \( W_{al} \) is the matrix of the connection weights between neurons; \( \alpha_y, \beta \) is the time constant that connects the input and variable threshold functions; \( V_y, \alpha_y \) is the amplitude coefficient of the connection input and variable threshold function; \( \alpha \) is the link strength factor. If \( U_y(n) \geq \theta_y(n) \), the neurons produce a pulse, called an ignition. PCNN model structure shown in Figure 3.

![Figure 3: Simplified model of PCNN](image)

Adaptive PCNN:

In traditional PCNN model, all neuronal link strength \( \beta \) is a constant value, but the visual characteristics of the human eye are not consistent with all characteristics of the image, and the neuron connection strength \( \beta \) is not fixed but varies according to the characteristics of the image at any time. Since the pixel values of image pixels are constantly changing, each neuron should have its own link strength according to the regional characteristics of the image. In addition, since the link strength \( \beta \) is in the PCNN model Therefore, the link strength \( \beta \) also reflects the regional characteristics of the image. In order to obtain a better fusion effect by using the PCNN model, this paper chooses to improve the Laplacian energy and IEOl as neuronal link strength \( \beta \), improve Laplacian energy and can effectively represent the image edge, direction and other details, suitable for high-frequency subband fusion, to improve the Laplacian energy and is defined as follows:

\[ ML_{k,j}(i, j) = \left| 2d_{k,j}(i, j) - d_{k,j}(i-1, j) - d_{k,j}(i+1, j) \right| + \left| 2d_{k,j}(i, j) - d_{k,j}(i, j-1) - d_{k,j}(i, j+1) \right| \]

where, \( d_{k,j}(i, j) \) denotes the coefficient of the high frequency subband in layer \( K \), direction \( L \), and the weighted Laplacian energy in the neighborhood at \((i, j)\) is:

\[ IEO_L_{k,j}(i, j) = \sum_{i-1}^{i+1} \sum_{j-1}^{j+1} w(i, j) \times ML_{k,j}(i, j, j) \]

Where \( w \) is the weight matrix,

\[
\begin{bmatrix}
1 & 2 & 1 \\
2 & 4 & 2 \\
1 & 2 & 1
\end{bmatrix}
\]

www.ijeijournal.com Page | 45
3.3.3 High frequency sub-band fusion rule

The source image is decomposed by NSCT to obtain three layers of high frequency sub-bands. The PCNN method is used for each layer to obtain the firing times of each pixel point of each layer of high frequency sub-bands. According to the number of ignition times to fuse, the fusion rule is as follows:

\[
d^f_{k,l}(i,j) = \begin{cases} 
  d^m_{k,l}(i,j), & \text{if } R^m_{k,l} > R^f_{k,l} \\
  d^n_{k,l}(i,j), & \text{if } R^n_{k,l} < R^f_{k,l} \\
  (d^m_{k,l}(i,j) + d^n_{k,l}(i,j))/2, & \text{otherwise}
\end{cases}
\]  

(16)

Where: \(d^f_{k,l}(i,j)\) is the high frequency subband fusion coefficient, \(d^m_{k,l}(i,j)\), \(d^n_{k,l}(i,j)\) respectively represents the coefficients of the high frequency subband M and N in layer \(K\), direction \(L\).

4 Experiment

In order to verify the effectiveness of this method, simulation experiments are carried out on the Matlab2012a platform. 30 groups of color fundus images and fluorescein fundus images collected from Tianjin Medical University of Second Hospital were used as experimental subjects. Each group of images are the same person with the same eye at the same time by different types of sensors shooting, there are differences issues in spatial location and resolution. So, before fusion, we need to unify the resolution of two images and register them. In this paper, we select 4 groups images of lesions of varying degrees to experiment. The fluorescence fundus images and the color fundus images are shown in Figure 4 to 7 (a),(b),(c). (a) is the fluorescence fundus image; (b) is the color fundus image before registration; (c) is the color fundus image after registration. because the fundus image are different from other medical fusion images, the general fusion method is not suitable for fundus image fusion. Therefore, two representative methods for fundus image fusion are selected for comparison. (d) is the fusion result of the literature [11] method, (e) is the fusion result of the literature [12] method, (f) is the fusion result of this paper.

Figure 4 The first set of images and the fusion of different methods
Figure 5 The second set of images and the fusion of different methods
From the fusion of the results of view, The fusion method of gray mean value used in [12] reduced the contrast of the whole image, the information of disc and the loss of lesion information were more. The fusion method in [13] and this paper method show the fusion results retain most of the color fundus image information, it can clearly show the bleeding point, exudate lesions, and increase the vascular information and micro-aneurysm lesions of fluorescence imaging fundus image, achieving the complementarity of the two types of images. In this
paper, the fusion of the images is superior to other representative methods in detail performance, such as the edge of lesion and microvascular.

In order to further analyze the fusion results, we select average gradient (AG) and information entropy (IE) to evaluate the above three representative methods, as shown in Table 1. As can be seen from the table, this method of paper is higher than other representative methods in the evaluation index, more information is obtained from the source image, and more clearly on the lesion display.

Table 1: Comparison of Objective Evaluation Indexes of Different Fusion Methods

<table>
<thead>
<tr>
<th>Image</th>
<th>Evaluation index</th>
<th>Fusion method</th>
</tr>
</thead>
<tbody>
<tr>
<td>first</td>
<td>AG</td>
<td>4.8184</td>
</tr>
<tr>
<td></td>
<td>IE</td>
<td>4.4144</td>
</tr>
<tr>
<td>second</td>
<td>AG</td>
<td>3.6210</td>
</tr>
<tr>
<td></td>
<td>IE</td>
<td>4.6541</td>
</tr>
<tr>
<td>third</td>
<td>AG</td>
<td>2.5579</td>
</tr>
<tr>
<td></td>
<td>IE</td>
<td>4.8784</td>
</tr>
<tr>
<td>fourth</td>
<td>AG</td>
<td>1.7970</td>
</tr>
<tr>
<td></td>
<td>IE</td>
<td>3.9590</td>
</tr>
</tbody>
</table>

IV. CONCLUSION

The fusion of fundus images requires that the fused image as accurately as possible can reflect the lesion, blood vessels and other information. In this paper, a method for registration and fusion of fundus images based on NSCT and adaptive PCNN is proposed: The color fundus image and the fluorescence contrast fundus image were NSCT decomposed in each channel of R, G, B to obtain the high-frequency sub-bands and low-frequency sub-band of the two images and the fusion of the low-frequency sub-band with the regional energy. Which can effectively preserve the complementary information between the two kinds of fundus images. For the high-frequency part of the two images, a simplified PCNN model is used for processing. Considering that the link strength parameter of the PCNN model has a great influence on the fusion result, the Laplacian energy as PCNN adaptive link strength. By comparing the number of ignition times to be fused, finally the NSCT inverse transformation outputs the fusion result of the channel, and finally the three-channel fusion result is synthesized into the total fusion result. The experimental results show that the method of this paper is superior to other methods in both subjective and objective evaluation.

REFERENCE