Image Fusion Based on Nonsampled Contourlet Transform

Ping Li

Laboratory Center, Hunan International Economics University, Changsha 410205, Hunan, China

Abstract

In order to make full use of the image information and reduce storage space needed by a lot of image, the image fusion technology arises at such historic moment. Image fusion technology can eliminate data redundancy between multi-sensors, improve the utilization rate of information, obtain more accurate and more abundant information than any single data, generate a synthetic image with new spatial, spectral and time characteristics, make the fusion image more conducive to the analysis and understanding, and more clear on the vision. Contourlet transform is a new two-dimensional image expression method with multi-resolution, multi-direction, time-frequency local and anisotropic characteristics, which has a broad application prospect in the image processing field. But contourlet transform does not have the translation invariant feature, and its application effect in some image processing fields is not very ideal. Nonsampled contourlet transform (NSCT) effectively overcomes the defect that the contourlet transform does not have the translation invariant feature. This paper proposes an image fusion algorithm based on NSCT. Such algorithm conducts NSCT decomposition on different source images, and different sub-bands adopt different fusion rules, and finally fusion image will be finally obtained by conducting NSCT inverse transform on coefficient fused. The experimental results show that, when compared with other fusion algorithms, this algorithm in this paper can improve the quality of fusion and facilitate the further use of the image.

Key words: Image Fusion, Contourlet Transform, Objective Evaluation.

1. INTRODUCTION

Under some circumstances, single image information is not sufficient for the detection, analysis and understanding of the targets or scenes due to the impact of such factors as the target state, the lighting, the environmental conditions, the target position and the inherent characteristics of the sensors, instead, it requires the combination of multiple images in order to obtain more comprehensive information(Salman and Lars, 2012). By fusing multiple sensor images, image fusion can overcome the limitations and differences of single sensor image in geometry, spectrum and spatial resolution, improve the image quality, facilitate the positioning, recognition and interpretation of physical phenomena and events and help to understand the image and obtain the interesting information. For the past few years, together with the development of wavelet theory and its applications, the inherent characteristics and the perfect reconstruction ability of wavelet transform has made sure that there is no information loss and redundant information in the signal decomposition and decomposed the image into average image and detailed image, which represent different image structures respectively; therefore, it is quite easy to extract the structural information and detail information of the original image. Wavelet analysis provides selective image coincident with human visual system(Jie and Hui, 2016; David and Franz, 2013).

Image fusion is first used in remote sensing image analysis. In 1979, Dalry recombines radar and satellite image and uses it in geological interpretation; in the 1980s, it is gradually applied in general image processing, in the 1990s, it becomes a research focus with the application fields of remote sensing, optical image processing and medical image processing and in the past years, it has become a research hotspot in computer vision, automatic target recognition, imaging navigation and guidance(Aslantas and Bendes, 2015). The early typical methods include ratio weighted method, Bovey transform method, principal component analysis (PCA) transform method and HIS transform method. As simple image fusion methods, these algorithms directly fuse in the pixel level, but they can’t perform transform decomposition on the image. These traditional image fusion methods do not conduct decomposition transform on source images. Afterwards, scientists have come up with the image fusion methods based on pyramid decomposition, including Laplacian pyramid, gradient pyramid and contrast pyramid. However, as there is correlation of decomposition component between the layers, the fusion result is not very ideal. Most of the traditional image fusion based on wavelet transform uses pixel average method and compared with the clear-cut regions in the original image, its corresponding regions in the fused image will be degraded; however, in contrast to the fuzzy regions, its corresponding regions will be improved. This method has reduced the image contrast in a certain extent and the result is not very good. At present, the image fusion methods in the wavelet transform domain are based on the fusion operators such as the maximum...
value, the local energy and the local variance. The proposal and development of multi-scale geometric analysis theory (MGA) makes up for such defects of the wavelet transform (Amina and Abdul, et al., 2015). In 2002, Do and Vetterli put forward a new non-self-adaptive direction multi-scale analysis method- contourlet transform, on the basis of wavelet multi-scale analysis thought. Contourlet transform can realize the decomposition on any direction and any scale, is skilled in describing the image contour and directional texture information, thus well making up for the inadequacy of the wavelet transform (Changdong and Zhigang et al., 2014; Luyi and Changming et al., 2015).

Firstly, this paper analyzes and discusses the basic theory of image fusion, including the process and the levels of image fusion as well as the common methods of pixel-level image fusion. Then, it makes a simple introduction of wavelet theory, describes the fundamental principles of wavelet theory and analyzes the theoretical framework, observation and reconstruction of wavelet decomposition. Since non-subsampled Contourlet transform has translation invariance and it can accurately capture the edge contour information and the texture detail information of the image, this paper proposes an image fusion method based on non-subsampled Contourlet transform. Finally, it offers the evaluation, the experimental simulation and analysis of image fusion.

2. IMAGE FUSION

Image fusion combines two or more images into a new image through a specific algorithm. By extracting and synthesizing multiple image information, it obtains a more accurate, more comprehensive and more reliable image description of the same target. Fusion algorithm shall make full use of the complementary information of various source images, making the fused image more suitable for human vision. According to the stage of the fusion in the processing and the degree of information abstraction, multi-source image fusion can be divided into three levels: pixel-level fusion, feature-level fusion and decision-level fusion. Among them, pixel-level fusion is not only the lowest level of fusion, but also the foundation of the other two levels. It performs fusion processing on the corresponding pixels in the source images and preserves as much image information as possible with higher accuracy. Pixel-level fusion directly makes information integration processing on the image pixels so as to produce a clearer image which contains more information. Normally, it requires accurate registration of the original images in the space. If the images have different resolutions, mapping processing is required before fusion (Yifang, 2016). Pixel-level image fusion refers to the information fusion in the level of basic data. It directly conducts fusion on the gray levels of the source images according to certain fusion rules. Its main task is to directly perform fusion processing on such information as the targets and the background of multi-source images and its ability to preserve information is stronger than those of the decision-level fusion and feature-level fusion, however, it has higher requirements on the registration accuracy. The process of pixel-level fusion includes four steps: pre-processing, image registration, fusion and inverse transformation (image reconstruction), at present, most research focuses on this level (Mansour and Shadrokh, et al., 2015; Ming and Wei, et al., 2013). The procedures of pixel-level fusion are shown in Figure 1.
recognition. Image feature matching has strong adaptability on the impact of such factors as change of view, affine transformation and noises.

To start with, it extracts features from the image to be registered. Then, it uses the similarity measurement and some constraint conditions to determine geometric transformation. Finally, it applies the above transformation in the image to be matched. The common features used in the matching include edge, contour, straight line, point of interest, color and texture, as indicated in Fig.2.

![Figure 2. Matching based on features](image)

Feature space is required to be determined before the matching method based on feature extracts the selected features from the image in the pre-processing. Normally, the selection of feature space shall make sure that the features of the reference image are of the same type as those of the image to be registered with certain invariance. If the two images to be registered have different imaging time or come from the sensors with different physical parameters, the features extracted will not change dramatically and they will remain the same in the transforms like rotation, translation and scale of two images. The selected matching features of the overlapping regions of two images shall be distributed as evenly as possible. Besides, the selected features shall be extracted easily (Chun-Man and Bao-Long, et al, 2014).

3. DISCRETIZATION OF CONTINUOUS WAVELET AND NONSUBSAMPLED CONTOURLET TRANSFORM

3.1. Discretization of Continuous Wavelet

Due to the redundancy of the continuous wavelet transform, in order to reconstruct the signal, it is necessary to carry out certain discretization towards the transform domain’s variable $a$ and $b$, so as to eliminate the redundancy in transform. Actually, when the value is $b = \frac{k}{2^j}, a = \frac{1}{2^j}; j,k \in \mathbb{Z}$, then

$$\psi_{a,b}(t) = \psi_{\frac{1}{2^j}, \frac{k}{2^j}}(t) = 2^{j/2} \psi(2^j t - k) \quad (1)$$

Usually shorten as $\psi_{j,k}(t)$, the transform form is

$$WT\left(\frac{1}{2^j}, \frac{k}{2^j}\right) = \{f, \psi_{j,k}\} \quad (2)$$

In order to reconstruct signal $f(t)$, it is required that $\left\{\psi_{j,k}\right\}_{j,k \in \mathbb{Z}}$ is $L^2(R)$’s Riesz base.

The pic.3 below is one-dimension discrete wavelet decomposition.

![Figure 3. One-dimension discrete wavelet decomposition](image)
First, respectively carry out the wavelet transform on two (or more) images that have been registered to be decomposed into wavelet coefficient, and low-frequency coefficient reflects the basic outline of the image, while the high-frequency one reflects just details. The larger the image pixel is, the greater its impedance function towards the multi-scale decomposition is. For the clarity, the high-resolution image distortion effect is not obvious as the decomposition scale increases. Then, fuse the corresponding wavelet coefficients according to certain rules. Finally, carry out the inverse transform on fused coefficients, so as to realize image reconstruction, and then the fused image will be obtained (Anoop and Mathew, et al., 2014; Rajiv and Khare, 2014).

3.2. Realization Process of Image Fusion Based on Nonsubsampled Contourlet Transform Multi-Direction Decomposition

Multi-direction decomposition and base function support interval of contourlet transform have the "rectangular" structure changing as the scale’s length-width ratio, and thus making contourlet transform effectively capture the geometrical structure features of the image information. More importantly, contourlet transform combines the image’s multi-scale and multi-direction in a flexible and organic way, and thus it can describe the image accurately and optimally. Suppose there are registered source image A and B to be fuse, and fused image F, then, the fusion algorithm process steps are as follows:

1. First of all, adopt Laplacian Pyramid transform (LP) to conduct the multi-scale decomposition on the image to capture the singular point, then, realize multi-direction decomposition on the band-pass components decomposed by each level of the pyramid with directional filter bank (DFB) to synthetize singular “points” in the same direction into “line”, thus, sub-band images in different directions can be obtained after band-pass images obtained after the decomposition by LP transfers to DFB, and through iteration, contourlet transform can decompose the image into sub-band images in multiple scales and multiple directions. Sub-band coefficients in multiple scales and multiple directions are obtained after NSCT multi-scale decomposition is operated on the source image. These sub-band coefficients are:

\[
\{ \lambda^A_{j,k} (j \geq j_0) \} \quad \text{and} \quad \{ \lambda^B_{j,k} (j \geq j_0) \},
\]

in which, \( \lambda^A_{j,k} \) and \( \lambda^B_{j,k} \) are respectively the low-frequency sub-band coefficient matrix of image A and B, \( \lambda^A_{j,k} \) and \( \lambda^B_{j,k} \) are respectively sub-band coefficient matrix in number \( K \) direction, scale \( J \) of A and B.

2. Contourlet adopts a new DFB which is based on the DFB’s fan-shaped filter. Such new DFB can need not to adjust the input image, and has a simple expansion decomposition tree rule. Intuitively, with the appropriate combination of a DFB fan-shaped direction frequency segmenting filter and re-sampling shearing operation, DFB’s wedge frequency segmentation can be realized. Low-frequency sub-band adopts “coefficient of average” rule to fuse, namely:

\[
\lambda^F_{j,k} = \frac{1}{2} (\lambda^A_{j,k} + \lambda^B_{j,k}).
\]

3. DFB is applied to each level of high-frequency component obtained from LP decomposition, and 2’ n-th power objective direction sub-band in any scale can be obtained by the decomposition. Input high-pass sub-band of the image generated by LP sub-band decomposition each time to DFB to gradually connect singular points into linear structure, thereby capturing the contour in the image. Sub-band direction band-pass adopts the rule “based on maximum local energy and matching degree detection” to fuse, namely.

Carry out frontier continuation towards \( \lambda^F_{j,k} \) and \( \lambda^B_{j,k} \), then 8 neighborhood local energy \( E^F_{j,k} (x,y) \) and \( E^B_{j,k} (x,y) \) of the source image in the corresponding direction and corresponding resolution with \( (x,y) \) as the center are respectively:

\[
E^F_{j,k} (x,y) = \sum_{n=1}^{1} \sum_{m=1}^{1} \left[ \lambda^F_{j,k} (x+m,y+n) \right]^2, \quad E^B_{j,k} (x,y) = \sum_{n=1}^{1} \sum_{m=1}^{1} \left[ \lambda^B_{j,k} (x+m,y+n) \right]^2
\]

If \( E^F_{j,k} (x,y) \geq E^B_{j,k} (x,y) \), then \( \lambda^F_{j,k} (x,y) = \lambda^F_{j,k} (x,y) \), otherwise, \( \lambda^F_{j,k} (x,y) = \lambda^B_{j,k} (x,y) \).

In order to further improve the effect of the fusion rules based on area, using the local matching degree detection operator, define the direction of the source image corresponding, the corresponding resolution of local energy \( E^F_{j,k} (x,y) \) and \( E^B_{j,k} (x,y) \).

\[
M^B_{j,k} (x,y) = \frac{2E^F_{j,k} (x,y) \times E^B_{j,k} (x,y)}{\left[ E^F_{j,k} (x,y) \right]^2 + \left[ E^B_{j,k} (x,y) \right]^2}
\]
LP decomposition generates a low-pass sampling approach and a difference image between the original image and low-pass predication forecast. Further decompose the obtained low-pass images to gain low-pass image and difference image of the next layer, and in this way, multi-resolution decomposition of the image is obtained by gradual filtering. The final fused image \( F \) is obtained after the reconstruction by the inverse transform towards the fusion coefficient \( \{ \lambda_{f,j}, \lambda_{f,k} \} \). In practical application, the direction number generally increases along with the increased scale. In high dimensional case, each layer of LP decomposition only produces a band-pass image, because the LP filter bank only conduct sub-sampling on low-pass image so to avoid the frequency interference phenomenon (Ying and Jie, et al, 2014; Celle, 1969).

4. EXPERIMENTAL SIMULATION AND ANALYSIS

4.1. Objective Evaluation Parameters of Image Fusion

The objective evaluation of image fusion quality has four commonly used indexes which are respectively the information entropy, root mean square error, signal-to-noise ratio and average gradient.

(1) Information entropy

The information amount of source image changes before and after the fusion, so, the calculation of the information entropy can objectively measure the effect of image fusion quality. According to the Shannon information theory, the image information entropy of a pair of 2\(^n\)-th power level can be represented by Formula 5:

\[
E = -\sum_{k=0}^{2^n-1} P_k \log P_k
\]

In which, shows the image pixel–total pixels ratio when the gray scale of the total gray scale series \( P_k \) is \( k \). The greater the information entropy value is, the greater the amount of information contained is and the better the image fusion effect is.

(2) Root mean square error,

According to the theory of image fusion, we define the root mean square error RMSE between the fusion image and standard reference image as Formula 6:

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{H} \sum_{j=1}^{W} (X(i,j) - Y(i,j))^2}{HW}}
\]

In which, \( H \) and \( W \) respectively show the number of rows and columns of images. The smaller the root mean square value is, the better the fusion effect is.

(3) Signal to noise ratio

It is the parameter to reflect whether the image frame is clean without noise in terms of image quality. Peak signal-to-noise ratio (PSNR) is defined as the Formula 7:

\[
PSNR = 10 \log_2 \left( \frac{2^n - 1}{RMSE^2} \right)
\]

The higher the PSNR value is, the better the fusion effect is.

(4) Average gradient

The average gradient (grad) reflects the contrast of image details, which can be used as the standard to test the image resolution. The Formula 8 is:

\[
grad = \frac{1}{(H-1)(W-1)} \sum_{x=1}^{H-1} \sum_{y=1}^{W-1} \frac{1}{2} \left( \frac{\partial f(x,y)}{\partial x} \right)^2 + \left( \frac{\partial f(x,y)}{\partial y} \right)^2
\]

In which, \( f(x,y) \) shows the fusion image. The larger the gradient value is, the clearer the fusion image is.

4.2. Objective Evaluation of the Fusion Effect

We can evaluate the fusion effect of the algorithm by comparing these four parameters as the information entropy, the root mean square error, signal-to-noise ratio and average gradient. Objective parameters of fusion effect are respectively calculated by using common three image fusion methods with the matlab software and such method in this paper, after the comparison, the results of are as shown in Table 1,2, and the fused subjective effect is shown in Figure 4, 5.
(a) Image tartan to be fused  (b) Image sinsin to be fused

(c) Fusion image by wavelet transform  (d) Fusion image by NSCT

**Figure 4.** Image fusion effect 1

(a) Image woman to be fused  (b) Image wbarb to be fused

(c) Fusion image by wavelet transform  (d) Fusion image by NSCT

**Figure 5.** Image fusion effect 2
Comparison of different algorithms of the fusion effect using visual comparison and objective analysis. Visual comparison method is to rely on the human eye for the subjective evaluation of the fused image. In the case of large size, from the visual effect, the subjective evaluation fused effect of NSCT method is better than wavelet transform method, the whole image is more clear, edge and texture features more obvious, better to retain the source image information. Information entropy represents the average information of the image, the greater the information entropy, the more the average amount of information contained in the image, the better the quality of image fusion. The average gradient of an image is the evaluation index of the image clarity. The average gradient is larger, the image is more clear. We can see from objective evaluation indexes of the table 1,2 that, the larger the information entropy is, the greater the average gradient is, the greater the signal-to-noise ratio is, the smaller the root mean square error is, the better the fusion effect is. Therefore, it can conclude that the image fusion effect by the NSCT method in this paper is better.

5. CONCLUSIONS

Contourlet transform has the multi-resolution, local positioning, multi-directivity, neighboring boundary sampling and anisotropic properties, and its base function is distributed in multi-scale and multi-direction, and a small amount of coefficient can effectively capture image’s edge contour. This paper studies the principle of non-sampled contourlet transform (NSCT), and advantages of its multi-scale, localization, directionality and anisotropy, to put forward an image fusion method based on NSCT which can more effectively preserve the useful information and detail characteristics of the source image. At the end, with complete experiment and multiple data results, this paper proves the feasibility and effectiveness of the method proposed.

REFERENCES


