Infrared Physics & Technology 63 (2014) 110–118

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

Technique for gray-scale visual light and infrared image fusion based on non-subsampled shearlet transform

Weiwei Kong *

Department of Information Engineering, Engineering University of Armed Police Force, Xi'an, Shaanxi 710086, China

HIGHLIGHTS

- NSST is an excellent tool for multi-scale geometric analysis of images.
- Fusion rules including RAE and LDC are adopted to fuse sub-band images.
- The NSST–RAE–LDC algorithm is devised and given in detail.
- The NSST–RAE–LDC model is more effective compared with other classic ones.

ARTICLE INFO

Article history:
Received 25 September 2013
Available online 30 December 2013

Keywords:
Gray-scale visual light and infrared images
Non-subsampled shearlet transform
Region average energy
Local directional contrast

ABSTRACT

A novel image fusion technique based on NSST (non-subsampled shearlet transform) is presented, aiming at resolving the fusion problem of spatially gray-scale visual light and infrared images. NSST, as a new member of MGA (multi-scale geometric analysis) tools, possesses not only flexible direction features and optimal shift-invariance, but much better fusion performance and lower computational costs compared with several current popular MGA tools such as NSCT (non-subsampled contourlet transform). We specifically propose new rules for the fusion of low and high frequency sub-band coefficients of source images in the second step of the NSST-based image fusion algorithm. First, the source images are decomposed into different scales and directions using NSST. Then, the model of region average energy (RAE) is proposed and adopted to fuse the low frequency sub-band coefficients of the gray-scale visual light and infrared images. Third, the model of local directional contrast (LDC) is given and utilized to fuse the corresponding high frequency sub-band coefficients. Finally, the final fused image is obtained by using inverse NSST to all fused sub-images. In order to verify the effectiveness of the proposed technique, several current popular ones are compared over three different publicly available image sets using four evaluation metrics, and the experimental results demonstrate that the proposed technique performs better in both subjective and objective qualities.

1. Introduction

The purpose of image fusion is to integrate complementary and redundant information from multiple images of the same scene, which may come from different kinds of image sensors or the same one functioning in diverse modes, to create a single composite that contains all the typical features of the source images. The fused image will thus be more informative and suitable for human visual perception or for further image processing tasks. Especially, the fusion of images obtained from gray-scale visual light and infrared imaging sensors has played an increasingly significant role in enhancing the performance of some civilian and military applications, such as missile guidance, target recognition and security surveillance.

To date, a wide variety of techniques for image fusion have been designed and proposed. Basically, the main-streamed techniques can be categorized into two categories including spatial domain analysis (SDA) and multi-scale geometric analysis (MGA). Furthermore, SDA has also two main branches which are linear weighing (LW) and artificial neural networks (ANN), respectively. LW is commonly to set a series of weighted values in advance, aiming at deciding the respective participation rate of the pixels locating the same position in source images, and then operate on the pixels directly. Hence, we can often obtain the fused image easily. However, along with simplicity come several unavoidable side effects such as unbalanced contrast level and blurred image edges. Known as the third generation of ANN, pulse coupled neural network (PCNN) [1,2] is characterized by global coupling and pulse synchronization of neurons. In recent years, a series of PCNN-based fusion methods and their modified versions have been widely developed [3–7]. Large constructive work shows that PCNN is effective and

* Tel.: +86 15609231541.
E-mail address: kwwking@163.com

1350-4495/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.infrared.2013.12.016

applicable in the field of image fusion. Unfortunately, due to too many existing parameters and complicated function mechanism, PCNN-based methods nearly have a common limitation, namely the high computational complexity.

On the other hand, MGA-based image fusion is a biologically motivated method, which fuses multiple images at different spatial resolutions. Up to now, several methods based on MGA have been developed. The basic idea of such methods is to perform a multi-resolution decomposition on each source image, then integrate all these decompositions to form a composite representation, and finally reconstruct the fused image by performing an inverse multi-resolution transform. The common MGA approaches include the discrete wavelet transform, ridgelet transform, curvelet transform, contourlet transform, non-subsampled contourlet transform (NSCT), shearlet transform (ST), etc. Wavelet [8,9] is only good at dealing with the point-wise singularities. Ridgelet [10] is adept in capturing the information of straight lines, but its performance of dealing with curves is disappointing. Curvelet [11] does not have obvious property of spatial sampling. Contourlet transform [12,13] lacks the shift-invariance and causes pseudo-Gibbs phenomena around singularities. In contrast with the methods mentioned above, NSCT [14–16] and ST [17–21] have better fusion performance in terms of both subjective and objective qualities. However, the high computational cost of NSCT deteriorates the efficiency of image fusion seriously. The absence of shift-invariance in ST easily tends to the appearance of the 'Gibbs' phenomena. In order to overcome the above problems, the theory of non-subsampled shearlet transform (NSST) was proposed in [22], which not only absorbs the advantages of current popular MGA tools, but provides a superior mathematical structure and flexibility. Cao et al. [23] utilized NSST to complete the course of image fusion, but they did not analyze and discuss the fusion rules of sub-images further. Overall, there are few references on the application to image fusion using NSST. Therefore, further research on NSST is very necessary and it is hoped that NSST will be more competitive in image fusion.

Based on the analysis of previous methods mentioned above, it is necessary and meaningful to investigate the NSST further, and then try to apply it to the field of image fusion. Considering the increasing importance of the fusion of gray-scale visual light and infrared images, we propose a novel NSST-based image fusion method. We specifically present new fusion rules to merge high and low frequency sub-band coefficients in an appropriate way in order to achieve the best quality in the fused image. To evaluate the new method with several current popular fusion methods, simulation experiments are conducted from the angles of both fusion performance and computational efficiency.

The outline of this paper is as follows. In Section 2, the mathematical theory of NSST is briefly reviewed. Section 3 concretely describes our proposed algorithm for gray-scale visual light and infrared image fusion based on NSST in detail. Experimental results and performance analysis are presented and discussed in Section 4. Concluding remarks and future work are given in Section 5.

2. Non-subsampled shearlet transform

Let dimension $n = 2$, the affine systems of ST can be expressed as follows:

$$\{\psi_{j,l,k}(x) = |\det A|^{j/2}\psi(S^l A^j x - k) : l,\ j \in Z,\ k \in Z^2\} \quad (1)$$

where ψ is a collection of basis function and satisfies $\psi \in L^2(R^2)$, A denotes the anisotropy matrix for multi-scale partitions, S is a shear matrix for directional analysis. j, l, k are scale, direction and shift parameter respectively. A, S are both 2×2 invertible matrices and $|\det S| = 1$. For each $a > 0$ and $s \in \mathbf{R}$, the matrices of A and S are given as follows.

$$A = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}, \quad S = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} \quad (2)$$

The two matrices above play important roles in the process of shearlet transform. The former dominates the scaling of shearlet, and the latter only controls the orientation of shearlet.

Assume $a = 4$, $s = 1$, Eq. (2) can be modified.

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

For any $\xi = (\xi_1, \xi_2) \in \hat{R}^2, \xi_1 \neq 0$, the mathematical expression of basic function let $\psi^{(0)}$ for ST can be given as follows:

$$\hat{\psi}^{(0)}(\xi) = \hat{\psi}^{(0)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2(\xi_2/\xi_1) \quad (3)$$

where $\hat{\psi}$ is the Fourier transform of ψ. $\hat{\psi}_1 \in C^\infty(R)$, $\hat{\psi}_2 \in C^\infty(R)$ are both wavelet, and $\sup p\hat{\psi}_1 \subset [-1/2, -1/16] \cup [1/16, 1/2]$, $\sup p\hat{\psi}_2 \subset [-1, 1]$. It implies that $\hat{\psi}_0 \in C^\infty(R)$ and compactly supported with $\sup p\hat{\psi}_0 \subset [-1/2, 1/2]^2$. In addition, we assume that

$$\sum_{j \geqslant 0} |\hat{\psi}_1(2^{-2j}\omega)|^2 = 1, \quad |\omega| \geqslant 1/8 \quad (4)$$

And for each $j \geqslant 0$, $\hat{\psi}_2$ satisfies that

$$\sum_{l=-2^j}^{2^j-1} |\hat{\psi}_2(2^j\omega - l)| = 1, \quad |\omega| \leqslant 1 \quad (5)$$

From the conditions on the support of $\hat{\psi}_1, \hat{\psi}_2$ one can obtain that the function $\psi_{j,l,k}$ has the frequency support listed below:

$$\sup p\hat{\psi}^0_{j,l,k} \subset \{(\xi_1, \xi_2)|\xi_1 \in [-2^{2j-1}, -2^{2j-4}] \cup [2^{2j-4}, 2^{2j-1}], |\xi_2/\xi_1 + l2^{-j}| \leqslant 2^{-j}\} \quad (6)$$

That is, each element $\hat{\psi}_{j,l,k}$ is supported on a pair of trapeziform zones, whose sizes all approximate to $2^{2j} \times 2^j$. The tiling of the frequency by shearlet and the size of the frequency support of $\psi_{j,l,k}$ are illustrated in Fig. 1. Note that Fig. 1(b) only shows the frequency support for $\xi_1 > 0$; the figure of the other support for $\xi_1 < 0$ is symmetrical.

NSST combines the non-subsampled Laplacian pyramid (NSLP) transform with several different combinations of the shearing filters (SF). It is commonly acknowledged that NSST is the shift-invariant version of ST essentially. In order to eliminate the courses of up-sampling and sub-sampling, NSST utilizes NSLP filters as a substitute for the Laplacian pyramid filters used in the ST mechanism, so that it has superior performance in terms of shift-invariance, multi-scale and multi-directional properties. The discretization process of NSST is composed of two phases including multi-scale factorization and multi-directional factorization. NSLP is utilized to complete multi-scale factorization. The first phase ensures the multi-scale property by using two-channel non-subsampled filter bank, and one low frequency image and one high frequency image can be produced at each NSLP decomposition level. The subsequent NSLP decompositions are implemented to decompose the low frequency component available iteratively to capture the singularities in the image. The multi-directional factorization in NSST is realized via improved SF. These filters are formed by avoiding the sub-sampling to satisfy the property of shift-invariance. SF allows the direction decomposition with l stages in high frequency images from NSLP at each level and produces 2^l directional sub-images with the same size as the source image. Fig. 2 illustrates the two-level NSST decomposition of an image. The number of shearing directions is chosen to be 4 and 4 from finer to coarser scale.