Neurocomputing 182 (2016) 1–9

Contents lists available at ScienceDirect

Neurocomputing

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

Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion

Biting Yu [a], Bo Jia [a], Lu Ding [a], Zhengxiang Cai [a], Qi Wu [a,*], Rob Law [c], Jiayang Huang [b], Lei Song [a], Shan Fu [a]

[a] School of Electronic, Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China
[b] Shanghai Engineering Research Center of Civil Aircraft Health Monitoring, China
[c] School of Hotel and Tourism Management, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 4 December 2014
Received in revised form
15 September 2015
Accepted 8 October 2015
Communicated by Huaping Liu
Available online 31 October 2015

Keywords:
Multi-focus image fusion
Dual-tree complex wavelet transform
Support vector machine
Bacterial foraging optimization

ABSTRACT

This study proposed a new method for multi-focus image fusion using hybrid wavelet and classifier. The image fusion process was formulated as a two-class classification problem: in and out-of-focus classes. First, a six-dimensional feature vector was extracted using sub-bands of dual-tree complex wavelet transform (DT-CWT) coefficients from the source images, which were then projected by a trained two-class support vector machine (SVM) to the class labels. A bacterial foraging optimization algorithm (BFOA) was developed to obtain the optimal parameters of the SVM. The output of the classification system was used as a decision matrix for fusing high-frequency wavelet coefficients from multi-focus source images in different directions and decomposition levels of the DT-CWT. After the high and low-frequency coefficients of the source images were fused, the final fused image was obtained using the inverse DT-CWT. Several existing methods were compared with the proposed method. Experimental results showed that our presented method outperformed the existing methods, in visual effect and in objective evaluation.

1. Introduction

Image fusion refers to an image processing technique that produces a new and improved single image, known as the fused image. This paper is concerned with the problem of multi-sensor pixel-level image fusion. Through image fusion, we aim to produce reliable methods that represent the salient information obtained from different imaging sensors, and fuse these details into a synthetic image. Thereafter, image fusion becomes more applicable for human visual perception and computer processing.

Many important applications, such as digital imaging, medical imaging, remote sensing and machine vision, need image fusion techniques [1–5]. This paper has concentrated on multi-focus image fusion. Optical imaging cameras suffer from the problem of finite depth of field, which makes taking an image with all objects contained “in focus” impossible. The solution to this problem is the multi-focus image fusion technique. This technique can be useful, in digital camera design or in industrial inspection applications, where the need to visualize objects at very short distances makes the preservation of the depth-of-focus difficult [1,6].

Various methods implement image fusion for multi-focus images. The simplest fusion method in spatial domain simply takes the pixel-by-pixel average of the source images, which leads to reduced contrast [1]. Some more reasonable methods are proposed, such as fusing source images with divided blocks or segmented regions, but they suffer from blackness in the fused image [7–9]. This study has focused on the wavelet-based approach, which is a subset of multi-scale decomposition based methods. The ability of the wavelet transform to capture important features in a picture is the reason for selecting multi-scale decomposition-based methods as a tool for image fusion [10]. Beyond this reason, the main reason is the time–frequency analysis of wavelet transform, while in-focus pixels of an image contain major high-frequency information. Therefore, this ability of wavelet transform can be used to determine in-focus pixels.

Although DWT (Discrete Wavelet Transform) has been successfully used for image de-noising, it has shortcomings such as shift variance, aliasing and lack of directionality [11]. Real valued wavelet transforms do not provide any phase information [12]. Phase information describes the amplitude and the local behavior of a function, according to [13]. DT-CWT (Dual-tree complex wavelet transform) provides shift invariance and better directionality than real valued wavelet transforms [14]. DT-CWT is proposed to capture additional edge information [15]. The higher directionality and shift invariance

* Corresponding author.
E-mail address: wuqi7812@sjtu.edu.cn (Q. Wu).

http://dx.doi.org/10.1016/j.neucom.2015.10.084
0925-2312/

properties of DT-CWT make it suitable for image fusion [16]. In obtaining the satisfactory fusion results, multi-scale based methods include three stages: decomposition, coefficients fusion, and reconstruction [17]. By reviewing existing papers, we find that the integration of DT-CWT and kernel method classifier for image fusion has gained little attention.

Therefore, the paper proposed a new multi-focus image fusion method based on the integration of DT-CWT and support vector machine with bacterial foraging optimization. SVM is based on statistical learning theory and specializes for a smaller number of samples [18] that is suitable for distinguishing the features of DT-CWT in normal images. Specifically, feature-based fusion rules were presented to merge high- and low-frequency wavelet coefficients for the best quality in the fused image. The key step in the proposed image fusion method is the use of BFO-SVM in selecting DT-CWT coefficient features, and this issue has been investigated in this study. The inter-scale dependencies among wavelet coefficients in the DT-CWT sub-bands were proposed to obtain a reliable decision matrix. First, four-feature vectors were obtained using six directional sub-bands of DT-CWT in the first decomposition level of the source images. Then, these feature vectors were classified into two classes through an optimally trained SVM classifier [19], and the output, a decision matrix, selected high-frequency wavelet coefficients between two source images. The classifier output was also used to select low-frequency wavelet coefficients between the source images by down sampling.

This paper is organized as follows. In Section 2, the proposed fusion algorithm based on the BFO-SVM classifier and DT-CWT is presented. Section 3 shows various results and comparisons. Finally, Section 4 concludes with a brief summary.

2. Proposed image fusion method

In this section, we propose a novel multi-focus image fusion method using DT-CWT. Let us suppose that A and B represent the different source images of the same size. The block diagram of the proposed method is shown in Fig. 1. The essential steps of the proposed image fusion are arranged as follows:

Step 1: j-Level DT-CWT decomposition of source images is performed, and six directional high-frequency coefficients are obtained as $x^j (j = 1, 2, \ldots, J)$ and low-frequency coefficients y^j.
Step 2: Four-feature vectors LF of high-frequency coefficients are extracted.
Step 3: Four-feature vectors are classified into two classes (in- and out-of-focus) through the BFO-SVM classification system, generating decision matrices DM.
Step 4: The high-frequency coefficients of images A and B are fused using a fusion rule presented in Section 2.4.1.
Step 5: The decision matrices DM are down-sampled from high-frequency coefficients to fit the size of low-frequency coefficients, and the low-frequency coefficients are fused by the corresponding fusion rule presented in Section 2.4.2.
Step 6: The final fused image F is reconstructed using the new high and low frequency coefficients by inverse DT-CWT.

2.1. DT-CWT

The ordinary DWT is shift variant, and aliasing occurs because of down-sampling. As a result, a small shift in the input signal can cause a very different set of wavelet coefficients produced at the output. To overcome this problem, Selesnick et al. [14] and Kingsbury [20] introduced a new kind of wavelet transform DT-CW, which allowed perfect reconstruction while still providing the other advantages of complex wavelets, as illustrated in Fig. 2.

The DT-CWT yields perfect reconstruction by using two parallel decimated filter-bank trees with real-valued coefficients generated at each tree. At each scale, the DT-DWT produces six directional sub-bands d, oriented at $\pm 15°$, $\pm 45°$, and $\pm 75°$. The two-dimensional DT-CWT decomposes a 2-D image f to different scales x^j as:

$$x = \{x^1, x^2, \ldots, x^J, y^J\} \tag{1}$$

$$x^j = \begin{cases} x^j_{\mathrm{Re},1}(l,k), x^j_{\mathrm{Re},2}(l,k), \ldots, x^j_{\mathrm{Re},6}(l,k) \\ x^j_{\mathrm{Im},1}(l,k), x^j_{\mathrm{Im},2}(l,k), \ldots, x^j_{\mathrm{Im},6}(l,k) \end{cases} \tag{2}$$

where x^j is the high frequency sub-bands at level j, y^J is the low frequency sub-bands at the last decomposition level, x^j is composed of six directional sub-bands with real and imaginary parts, and (l, k) is the spatial position of the coefficients.

2.2. Feature extraction

Wavelet coefficient based features can reflect the image's detailed information well. Thereafter, some local features (i.e., mean, energy, standard deviation, and normalized Shannon entropy) from the DT-CWT coefficients were proposed to distinguish in and out-of-focus regions. A local window (3×3) was applied in calculating the local features at the six directions. Their

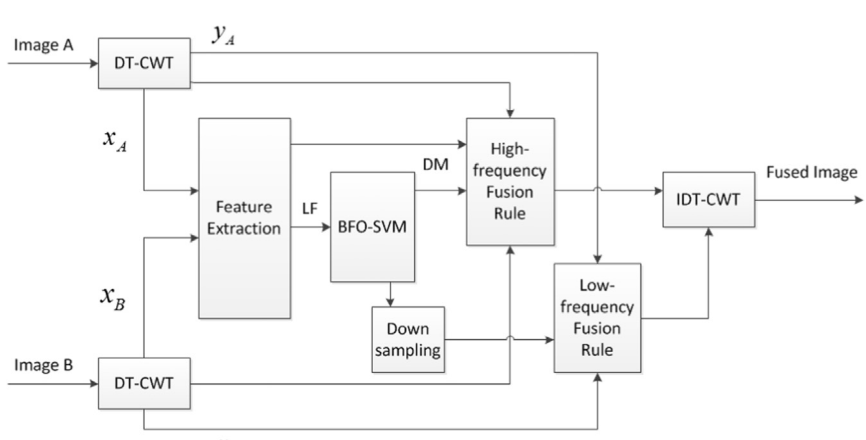

Fig. 1. Block diagram of the proposed method.