Infrared Physics & Technology 64 (2014) 134–142

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

A fusion method for visible and infrared images based on contrast pyramid with teaching learning based optimization

Haiyan Jin [a,b,*], Yanyan Wang [a]

[a] School of Computer Science & Engineering, Xi'an University of Technology, Xi'an 710048, China
[b] Shaanxi Key Laboratory for Network Computing and Security Technology, Xi'an 710048, China

HIGHLIGHTS

- This paper presents an optimized method for visible and infrared images fusion employing TLBO method.
- Fusion coefficients are non-linear adjusted adaptively by fitness function.
- Proposed method outperforms other methods in both visual effect and objective evaluation.

ARTICLE INFO

Article history:
Received 31 December 2013
Available online 13 March 2014

Keywords:
Image fusion
TLBO
Contrast pyramid
Optimization

ABSTRACT

This paper proposes a novel image fusion scheme based on contrast pyramid (CP) with teaching learning based optimization (TLBO) for visible and infrared images under different spectrum of complicated scene. Firstly, CP decomposition is employed into every level of each original image. Then, we introduce TLBO to optimizing fusion coefficients, which will be changed under teaching phase and learner phase of TLBO, so that the weighted coefficients can be automatically adjusted according to fitness function, namely the evaluation standards of image quality. At last, obtain fusion results by the inverse transformation of CP. Compared with existing methods, experimental results show that our method is effective and the fused images are more suitable for further human visual or machine perception.

1. Introduction

The spectrum is divided into IR, visible and UV spectroscopy, which have their own principles and characteristics of imaging, in different wavelength regions. In practical applications, multi-spectral resolution and high spatial resolution images are more widely used. Thus, the visible and infrared image fusion has become a hot topic in the field of image fusion. It is hoped that through some kind of image fusion techniques to fully complement both information, and access to more detail and accurate information on the scene, it is possible to locate the position of the heat source for identifying camouflage, testing resources, night vision and so on.

Image fusion refers to a multichannel source image data acquired on the same object after the image processing and computer technology, enabling to maximally extract the favorable information of each channel, and finally integrated into a high-quality image, the purpose is to improve the utilization of image information, the accuracy and reliability of the computer interprets and the spatial resolution and spectral resolution of the original image, which will help to monitor. This paper relates to the visible and infrared images fusion with having a wide range of practical value in many areas, such as electronics testing, military combat, lunar exploration, and medical image analysis.

Nowadays, there are already a lot of methods in the field of images fusion. For example, the intensity-hue-saturation (IHS) transposition, [1] principal component analysis (PCA), [2] Brovey algorithm, [3] wavelet-based methods, [4–7] or various pyramid algorithms such as contrast pyramid [8–11], etc. They are relatively simple image fusion method and most widely used, but there are still various deficiencies. For instance, the former three fusion methods can lead to different degrees of spectral distortion and lower spatial resolution. The wavelet-based fusion methods may generate "ringing" effects. CP-based fusion method can obtain better results, but its fusion coefficients are set by experience and are constants, there is very great blindness. We put forward a new method that TLBO [12–14] approaches to maximally optimize fusion coefficients on the basis of CP decomposition, changing the previous pattern of fixed coefficients selection, to achieve superior fusion effect.

* Corresponding author. Current address: P.O. Box 666, No. 5 South Jinhua Road, Xi'an 710048, China.

E-mail address: jinhaiyan@xaut.edu.cn (H. Jin).

http://dx.doi.org/10.1016/j.infrared.2014.02.013
1350-4495/

The rest of the paper is organized as follows. In Section 2, the decomposition and reconstruction of CP is briefly reviewed, and TLBO is simply introduced in Section 3. Then the application of TLBO for optimizing coefficients is described in detail in Section 4, Section 5 explains the principle and fitness function of the proposed method. The effect images and analysis of experiments are showed in Section 6. Finally, the conclusion is summarized in Section 7.

2. Decomposition and reconstruction of CP

The pyramid of an image is calculated from bottom to top. It is decomposed each image into pyramid image sequence of multiple scales, high resolution image in the lower layer, low resolution image in the upper, the size of the upper image is a quarter of previous layer image'. In consideration of CP derived from Gaussian pyramid, every level of which is formed by the previous one through filtering, we first build the Gaussian window function. Suppose w denote 5×5 window function, and satisfy the following constraints:

(1) Separability

$$w(m,n) = w(m)w(n), m \in [-2,2], n \in [-2,2] \tag{1}$$

(2) Normalized

$$\sum_{m=-2}^{2} w(m) = 1 \tag{2}$$

(3) Symmetry

$$w(m) = w(-m) \tag{3}$$

(4) Parity items such as contribution

$$w(-2) + w(2) + w(0) = w(-1) + w(1) \tag{4}$$

These constraints are used to ensure the performance of the low-pass, and the brightness of images is smoothness after shrinking and expanding, it does not appear the joint effect. We can construct: $w(0) = 3/8$, $w(1) = w(-1) = 1/4$, $w(2) = w(-2) = 1/16$, the window function can be mathematically expressed as:

$$w = \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \tag{5}$$

Gaussian pyramid as a set of image sequence, each level are made at the next higher level image through a low-pass filter and the down-sampling process. The frequency band of two adjacent is reduced by 1/8 rate, the image size by 1/4 rate decreases. Here we set the source image is G_0 (i, j), $i \leqslant M^0$, $j \leqslant N^0$, that is Gaussian pyramid's bottom image, M^0, N^0 is the number of columns and rows of the image respectively. G_l represents the first l-level of Gaussian pyramid, $0 < l \leqslant N$ (N is the total number of layers). To get G_l, it must firstly take G_{l-1} and the above constructed window function w convolve, and then the convolution result is done the down-sampling in interlaced columns and rows. Computation formulation is as follows.

$$G_l = \sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n)G_{l-1}(2i+m, 2j+n),\ 0 < l \leqslant N,$$
$$0 \leqslant i < u,\ 0 \leqslant j < v \tag{6}$$

where u, v, denote the number of columns and rows of the first l layer pyramid image, respectively; w is namely the above defined 5×5 low-pass filter.

After winning Gaussian pyramid, it will be interpolated to enlarge. We introduce enlarge operator, that is interpolate new samples between the given numerical by using interpolation method. A certain level of the pyramid image is extended into the size of the former level image. Thus, it can solve the problem of these levels' different in sample density. If carry out the following operation on G_l, we can get the new image owning the same size as G_{l-1}.

$$G^*_{l-1}(i,j) = 4\sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n)G'_l\left(\frac{i+m}{2},\frac{j+n}{2}\right), 0 < l \leqslant N,$$
$$0 \leqslant i < u,\ 0 \leqslant j < v \tag{7}$$

where

$$G'_l\left(\frac{i+m}{2},\frac{j+n}{2}\right) = \begin{cases} G_l\left(\frac{i+m}{2},\frac{j+n}{2}\right) & \text{where } (i+m/2, j+n/2) \text{ is integer coordinates} \\ 0 & \text{otherwise} \end{cases}$$

Every level of CP C_l, is a ratio of the corresponding Gaussian image at all levels. It is defined as

$$\left.\begin{array}{ll} C_l = G_l/G^*_l, & 0 \leqslant l \leqslant N-1 \\ C_N = G_N, & l = N \end{array}\right\} \tag{8}$$

Than we consider CP reconstruction, which is the inverse process of decomposition. According to the iteration of formulation (8), set $l = N$, $N-1$, ..., 0, starting from the top of CP and turning down, we can get each layer of Gaussian pyramid: G_n, G_{n-1}, ..., G_0. The reconstruction image G_0 is the final fusion image. The formulation is as follows.

$$\left.\begin{array}{ll} G_N = C_N, & l = N \\ G_l = C_l G^*_l, & 0 \leqslant l < N \end{array}\right\} \tag{9}$$

where G^*_l is the same size as G_l after G_{l+1} interpolated.

The original images are decomposed as a series of decomposition layers, which have different resolutions and spatial frequency, by the multiscale CP decomposition. At the same time, each of the decomposition levels reflects the contrast information of the corresponding spatial frequency. The fusion process is conducted in every spatial frequency layer respectively, so it may be aimed at the features and details of different frequency bands on different decomposition levels by using different fusion operators, for the purpose of highlight the features and details of the specific frequency band. It follows the principle that the human retinal image is processed in different frequency channel, so it may gain the fusion effect which is more close to the human visual characteristics. Through many literatures and simulation experiments, CP fusion method is better than general pyramidal decomposition and wavelet transform for visible and infrared images fusion. However, there are inherent weaknesses: the selection of weighted coefficients is mostly based on experience, which is the coefficients values are prior set, it can be easy to cause unstable fusion effect. To improve this situation, we adopt the heuristic evolutionary algorithm, TLBO, which can search and estimate the fusion coefficient in certain

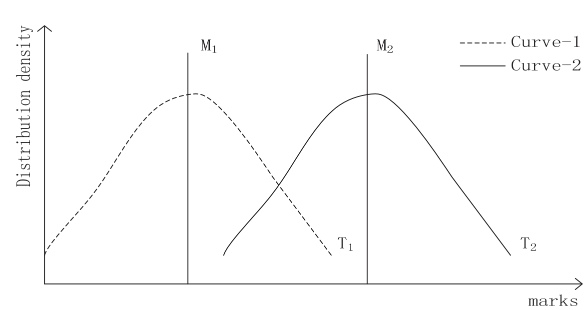

Fig. 1. Model for the distribution of marks for two groups.