J. Vis. Commun. Image R. 40 (2016) 218–224

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Image fusion based on visual salient features and the cross-contrast ☆

Jianhua Adu [a], Shenghua Xie [b,*], Jianhong Gan [a]

[a] Software Department, Chengdu University of Information Technology, Chengdu 610225, China
[b] Sichuan Academy of Medical Sciences & Sichuan Provincial People's Hospital, Sichuan Provincial Key Laboratory of Ultrasound in Cardiac Electrophysiology and of Biomechanics, Chengdu 610072, China

ARTICLE INFO

Article history:
Received 17 May 2016
Revised 20 June 2016
Accepted 25 June 2016
Available online 27 June 2016

Keywords:
Image fusion
Nonsubsampled contourlet transform
Visual salient features
The cross-contrast

ABSTRACT

To extract and combine the features of the original images, a novel algorithm based on visual salient features and the cross-contrast is proposed in this paper. Original images were decomposed into low frequency subband coefficients and bandpass direction subband coefficients by using the nonsubsampled contourlet transform. Three maps of visual salient features are constructed based on visual salient features the local energy, the contrast and the gradient respectively, and low-frequency subband coefficients are got by utilizing these visual saliency maps. The cross-contrast is obtained by computing the ratio between the local gray mean of bandpass direction subband coefficients and the local gray mean of fused low-frequency subband coefficients. Bandpass direction subband coefficients is goted by the cross-contrast. Comparison experiments have been performed on different image sets, and experimental results demonstrate that the proposed method performs better in both subjective and objective qualities.

1. Introduction

Image fusion is an active research area in optical signal processing, the objective of image fusion is to combine useful information from several images of the same picture or scene [1]. Therefore, multiple different images of one same scene may be acquired by different image sensors under different optic conditions or at different times to integrate different data so as to obtain more information [2]. Because the fused image contains the main features of several images which captured by different sensors, the target object in the same scene can be observed and distinguished more clearly, more comprehensively, more reliably. Now, as an important image analysis and computer vision technology, image fusion has widely applied to target recognition, computer vision, remote sensing, robot, medical image processing, military application, etc. Meanwhile, image fusion can provide more effective information for further computer image processing, such as high efficiency video processing, image classification, image segmentation, object recognition and detection [3–9].

In recent years, many effective image fusion methods have been proposed, such as the method based on Multi-scale transform (MST) [10], the method based on ICA or PCA [11], the method based on neural networks [12], the method based on SIFT [13] and the method based on morphological component [14]. Multi-scale transform (MST)-based fusion methods are the most popular and important tools in image processing, which are also effectively used for image fusion. There are many classical MST-based fusion methods such as pyramid-based ones, wavelet-based ones and multi-scale geometric analysis (MGA)-based ones. Pyramid-based ones include Laplacian pyramid (LP) [15,16], ratio of low-pass pyramid (RP) [17] and gradient pyramid (GP) [18,19]. The wavelet-based ones include discrete wavelet transform (DWT) [10,20], stationary wavelet transform (SWT) [21–24] and dual-tree complex wavelet transform(DTCWT) [25]. The multi-scale geometric analysis (MGA)-based ones include curvelet transform (CVT) [26,27], ridgelet transform [28], nonsubsampled contourlet transform (NSCT) [29–31] and nonsubsampled shearlet transformation(NSST) [32–34]. In general, the MST-based fusion methods consist of the following three steps [35,36]. First, the original images are decomposed into a multi-scale transform domain. Secondly, the transformed coefficients are merged with a given fusion rule. Finally, the fused image is reconstructed by performing the corresponding inverse transform over the merged coefficients. Therefore, it's obvious that the fusion rule of high-pass and low-pass subband image plays a crucial role for the result of image fusion. Moreover, transform domain also has a great impact on the fused results.

Human is primarily dependent on visual sense to obtain information from the outside world. The studies of human visual system (Human Visual system, HVS) have shown that during

☆ This paper has been recommended for acceptance by M.T. Sun.
* Corresponding author.
E-mail address: adjh_work@163.com (S. Xie).

http://dx.doi.org/10.1016/j.jvcir.2016.06.026
1047-3203/

observating and understanding a image, HVS is usually more concerned about salient features of the image [37–39]. Some analysis methods based on visual saliency also have been proposed to quickly detect salient area or targets in an image [40–43]. In this paper, three feature maps will be constructed based on visual saliency which are the local energy, the contrast and the gradient respectively, and low frequency subband coefficients are fused utilizing these visual feature salient maps. Then, a cross-contrast fusion method is used to get bandpass directional subband coefficients, and the cross-contrast represents the ratio between the local gray mean of the bandpass directional subband coefficients and the local gray mean of the fused low frequency subband coefficients. A comparative study of different MST-based methods is reported in [44], where Li et al. found that the NSCT-based method can generally achieve the best results. Therefore, in this paper, NSCT has been selected as MST-based fusion method. This paper is organized as follows. The following section briefly explains the principle of NSCT, and the Section 3 introduces image fusion based on nonsubsampled contourlet transform. the Section 4 introduces image fusion algorithm based on visual salient features and the cross-contrast. In Section 5, the results and analysis of experiments are presented. Finally, our conclusions are given in Section 6. Further, for brevity, in the subsequent part of this paper we use the abbreviation LFS and BDS, and define low frequency subband coefficients as LFS coefficients and bandpass directional subband coefficients as BDS coefficients

2. Non-subsampled contourlet transform

The tools of multiscale geometric analysis have been broadly used in image fusion. Nowadays, wavelet transform is an efficient tool to express the one-dimensional (1-D) piecewise smooth signals, but in the case of two-dimensional (2-D) signals, it cannot efficiently preserve edges of a nature image. In addition, separable wavelets are deficient in capturing only limited directional information and feature of multi-dimensional signals.

To overcome the drawbacks of wavelet in dealing with higher dimension signals, Do and Vetterli [45] recently pioneered a new representational system named contourlet, which is a real representation of 2-D signal. The contourlet transform uses the Laplacian pyramid (LP) for multi-scale decomposition, and the directional filter bank (DFB) for directional decomposition. The contourlet transform was proposed to address the lack of geometrical structure in the separable two-dimensional wavelet transform. Because of its filter bank structure, the contourlet transform is not shift-invariant. Afterwards, in 2006, a novel multi-scale decomposition method, the nonsubsampled contourlet transform evolving from contourlet transform, was proposed by da Cunha et al. [46]. The NSCT is not only with multi-scale, localization, and multi-direction, but also with properties of shift-invariance and the same size between each subband image and the original image. The NSCT not only retains the characteristics of contourlet, but also has other important properties of the shift invariance. The size of different subbands is identical, so it is easy to find the relationship among different subbands, which is beneficial for designing fusion rules. The contourlet transform and NSCT have a similar approach of decomposition and reconstruction. In NSCT, the multiscale analysis and the multidirection analysis are also separate, but both are shift-invariant. The construction of NSCT is based on a nonsubsampled pyramid filter banks (NSPFB) and nonsubsampled directional filter banks (NSDFB), and each subband image has the same size with the original image. Therefore, the NSCT is a flexible multi-scale, multi-direction, and shift-invariant image decomposition, as is displayed in Fig. 1. First, NSPFB is used to obtain a multiscale decomposition by using two-channel nonsubsampled 2-D filter banks. The NSPFB decomposition is similar to the 1-D nonsubsampled wavelet transform (NSWT) computed with the àtrous algorithm [47]. Second, NSDFB is used to split bandpass subband in each scale into different directions. Finally, we also can take an inverse transform to reconstruct an image by these coefficients obtained by NSCT. More details can be seen in [46]. Consequently, introduction of NSCT into image fusion could do justice to the good character of NSCT in effectively presenting features of original images. Although the work efficiency of NSCT is a bit slow, the results are excellent. Now, the hardware has a strong computing power, so it does not matter that the shortcoming of NSCT compared with its superior performance.

3. The image fusion based on nonsubsampled contourlet transform

Based on the above theory, NSCT can effectively be applied to image fusion. The image fusion based on nonsubsampled contourlet transform is usually done by the following steps [48].

3.1. Image decomposition

In this paper, the intensity of pixel (m, n) of a decomposed image can be defined as "image coefficient". Firstly, assume that there have been two original images f_1 and f_2 that are geometrically registered to each other. Secondly, original images f_1 and f_2 are separately decomposed into multiscale and multidirection with NSCT. Therefore, coefficients of two images $\{C_{i_0}^{f_1}(m,n), C_{i,l}^{f_1}(m,n)(i \geqslant i_0)\}$ and $\{C_{i_0}^{f_2}(m,n), C_{i,l}^{f_2}(m,n)(i \geqslant i_0)\}$ are obtained respectively. Here, $C_{i_0}(m,n)$ is obtained after a multiscale decomposition by using two-channel nonsubsampled 2-D filter banks, which denotes the lowpass subband coefficient of the pixel (m, n) at the coarsest scale, i.e., the i_0-th scale. $C_{i,l}(m, n)$ is obtained after splitting bandpass subband in each scale into different directions, which denotes the BDS coefficient of the pixel (m, n) at the i-th scale and the l-th direction. In this paper, $C_{i_0}(m,n)$ and $C_{i,l}(m, n)$ are obtained by NSCT tool provided by da Cunha et al. [46].

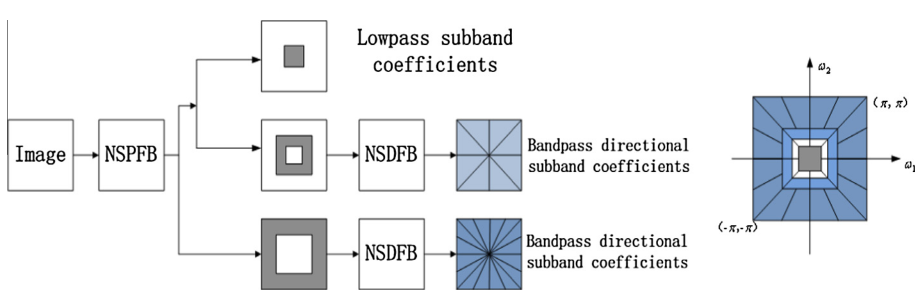

Fig. 1. Nonsubsampled contourlet transform decomposition framework.