Neurocomputing 157 (2015) 143–152

Contents lists available at ScienceDirect

Neurocomputing

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

A new contrast based multimodal medical image fusion framework

Gaurav Bhatnagar [a,*], Q.M. Jonathan Wu [b], Zheng Liu [c]

[a] Indian Institute of Technology Jodhpur, Jodhpur, India
[b] University of Windsor, Windsor, Canada
[c] Toyota Technological Institute, Nagoya, Japan

ARTICLE INFO

Article history:
Received 21 November 2013
Received in revised form
9 January 2015
Accepted 10 January 2015
Communicated by Huiyu Zhou
Available online 19 January 2015

Keywords:
Multimodal fusion
Medical imaging
Non-subsampled contourlet transform
Entropy
Directive contrast

ABSTRACT

Medical image fusion plays an important role in clinical applications such as image-guided surgery, image-guided radiotherapy, noninvasive diagnosis, and treatment planning. The main motivation is to fuse different multimodal information into a single output. In this instance, we propose a novel framework for spatially registered multimodal medical image fusion, which is primarily based on the non-subsampled contourlet transform (NSCT). The proposed method enables the decomposition of source medical images into low- and high-frequency bands in NSCT domain. Different fusion rules are then applied to the varied frequency bands of the transformed images. Fusion coefficients are achieved by the following fusion rule: low-frequency components are fused using an activity measure based on the normalized Shannon entropy, which essentially selects low-frequency components from the focused regions with high degree of clearness. In contrast, high-frequency components are fused using the directive contrast, which essentially collects all the informative textures from the source. Integrating these fusion rules, more spatial feature and functional information can be preserved and transferred into the fused images. The performance of the proposed framework is illustrated using four groups of human brain and two clinical bone images from different sources as our experimental subjects. The experimental results and comparison with other methods show the superior performance of the framework in both subjective and objective assessment criteria.

1. Introduction

To support more accurate clinical information to physicians for better diagnosis, multimodal medical images are needed, such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI), and magnetic resonance angiography (MRA). Medical image fusion helps physicians to extract features from different modalities that may not be normally visible in the images. For example, the CT image can show dense structures like bones and implants with less distortion, but it cannot detect physiological changes, while the MR image can provide normal and pathological soft tissues information, but it cannot support the bones information [1]. Even a single modality can provide complementary and occasionally conflicting information due to its dependence on variable parameters. For instance, T1 weighted MR imaging gives enhanced detail of anatomical structures whereas T2 weighted MR imaging gives greater contrast between normal and abnormal tissues. Therefore, only one kind of multimodal image may not be sufficient to provide accurate clinical requirements to the physicians [1].

So far, many image fusion frameworks have been proposed in the literature [2–9] with some specific for multimodal medical image fusion [10–21]. These frameworks can be broadly classified into three categories based on the stage at which the combination mechanism takes place. This characterization includes pixel-level or sensor-level, feature-level, and decision-level fusion [2]. Among these, the most popular framework is pixel-level fusion due to the advantage of containing the originally measured quantities, easy implementation and computationally efficient [7]. Hence, in this paper, we concentrate our efforts on pixel level-fusion, and the terms image fusion or fusion are intently used for pixel level fusion throughout the paper. The well-known pixel-level frameworks are based on principal component analysis (PCA), independent component analysis (ICA), gradient pyramid (GP) filtering, etc. [22–25]. These approaches are not fully suitable for the application of medical image fusion since the features are sensitive to the human visual system existing in different scales [12]. Therefore, a multi-scale or multiresolution analysis is more suitable for the fusion purposes. With the development of multiresolution analysis, wavelet transform has been identified as an ideal method for fusion. However, it is argued that wavelet decomposition is good at isolated discontinuities, but with a poor performance at edges

* Corresponding author.
E-mail addresses: goravb@iitj.ac.in (G. Bhatnagar), jwu@uwindsor.ca (Q.M.J. Wu), zheng.liu@ieee.org (Z. Liu).

http://dx.doi.org/10.1016/j.neucom.2015.01.025
0925-2312/

and textured regions. Further, it captures limited directional information along vertical, horizontal and diagonal directions [13]. These issues are rectified in a recent multiscale decomposition, namely contourlet and its non-subsampled version. Contourlet is a "true" 2-D sparse representation for 2-D signals like images where sparse expansion is expressed by contour segments. As a result, it can capture 2-D geometrical structures in visual information much more effectively than the traditional multiscale methods [26]. In contrast, NSCT inherits all the advantages of contourlet transform along with shift-invariance property and effectively suppressing pseudo-Gibbs phenomena. Hereafter, some authors have proposed image fusion framework using NSCT [17–21]. Among these, most of the frameworks are implemented in multi-focus fusion. If implemented for medical imaging, the results are not of the same quality as those for the multimodal medical image fusion. The main reason is the structure of medical images. Due to this fact, traditional fusion rules such as weighted average, absolute maximum, spatial frequency and saliency do not efficiently utilize prominent information present in the low- and high-frequency coefficients and result in the poor quality [21]. Therefore, two new fusion rules are proposed in this work to address these issues.

In this paper, a fully automated framework for medical image fusion is proposed in the non-subsampled contourlet transform (NSCT) domain. After the source images are decomposed by the NSCT, the coefficients of the low- and high-frequency portions are fused using two different fusion processes, which are chosen considering the physical meaning of the coefficients. Therefore, the coefficients of the low- and high-frequency bands are treated differently: the former is selected with an activity measurement process, and the latter is selected by a contrast based process. The fused image is then obtained by taking inverse NSCT transform on the fused low- and high-frequency coefficients. Both qualitative and quantitative performance evaluations are carried out to validate the proposed framework. The final fused images are obtained by applying inverse NSCT on the fused low- and high-frequency coefficients. Extensive experiments on different multi-modal CT/MRI and MR-T1/MR-T2 data-sets are carried out along with two clinical examples. Performance comparison of the proposed framework with the existing methods demonstrates the efficiency of the proposed method.

The rest of the paper is organized as follows. The NSCT is described in detail in Section 2 followed by the introduction of multimodal medical image fusion framework in Section 3. Experimental results and discussion are given in Section 4 and the concluding remarks are presented in Section 5.

2. Non-subsampled contourlet transform (NSCT)

NSCT based on the theory of countourlet transform (CT) is a kind of multi-scale and multi-direction computation framework of the discrete images [26]. It can be divided into two phases including non-subsampled pyramid (NSP) and non-subsampled directional filter bank (NSDFB). The former phase ensures the multiscale property by using two-channel non-subsampled filter bank, producing one low-frequency and one high-frequency image at each NSP decomposition level. Subsequent NSP decomposition stages are carried out to decompose the available low-frequency component iteratively to capture the singularities in the image. As a result, NSP results in $k+1$ sub-images, which consist of one low- and k high-frequency images having the same size as the source image where k denotes the number of decomposition levels. Fig. 1(a) shows the NSP decomposition with $k=3$ levels. The NSDFB is two-channel non-subsampled filter bank which is constructed by combining the directional fan filter banks. NSDFB allows the direction decomposition with l stages in high-frequency images from NSP at each scale and produces 2^l directional sub-images with the same size as the source image. Therefore, NSDFB offers the NSCT with the multi-direction property and provides us with more precise directional details information. A four channel NSDFB constructed with two-channel fan filter banks is illustrated in Fig. 1(b).

3. Proposed multimodal medical image fusion framework

The proposed framework realizes on a new definition of the directive contrast in NSCT domain, which takes a pair of source image denoted by A and B to generate a composite image F. The basic condition in the proposed framework is that all the source images must be registered in order to align the corresponding pixels. The definition of the directive contrast and the proposed fusion framework are described below.

3.1. Directive contrast in NSCT domain

The contrast feature measures the separation between the intensity values of a pixel and its neighboring pixels. The human visual system is highly sensitive to the intensity contrast rather than the intensity value itself. Generally, the same intensity value looks like a different intensity value depending on intensity values of neighboring pixels. Therefore, local contrast is developed and is defined as [27]

$$C=\frac{L-L_B}{L_B}=\frac{L_H}{L_B} \tag{1}$$

where L is the local luminance and L_B is the luminance of the local background. Generally, L_B is regarded as local low-frequency and hence, $L-L_B=L_H$ is treated as local high-frequency. This definition is further extended as directive contrast for the multimodal image fusion. These contrast extensions take high-frequency as the pixel value in multiresolution domain. However, considering single pixel is insufficient to determine whether the pixels are from clear parts or not. Therefore, directive contrast is integrated with the sum-modified Laplacian [28] to get salient features.

In general, the larger absolute values of high-frequency coefficients correspond to the sharper brightness in the image and lead to the salient features such as edges, lines, and region boundaries. However, these are very sensitive to the noise, which can be taken as the useful information and leading to misinterpretation of the actual information in the fused images. Hence, a proper way to select high-frequency coefficients is necessary to ensure better information interpretation. The sum-modified Laplacian is integrated with the directive contrast in NSCT domain to produce accurate salient features. Mathematically, the directive contrast in NSCT domain is given by

$$D_{l,\theta}(i,j)=\begin{cases}\dfrac{\mathrm{SML}_{l,\theta}(i,j)}{I_l(i,j)} & \text{if } I_l(i,j)\neq 0\\ \mathrm{SML}_{l,\theta}(i,j) & \text{if } I_l(i,j)=0\end{cases} \tag{2}$$

where $\mathrm{SML}_{l,\theta}$ is the sum-modified Laplacian of the NSCT frequency bands at scale l and orientation θ. On the other hand, $I_l(i,j)$ is the low-frequency sub-band at the coarsest level (l). The sum-modified Laplacian is defined by the following equation:

$$\mathrm{SML}_{l,\theta}(i,j)=\sum_{x=i-m}^{i+m}\sum_{y=j-n}^{j+n}\nabla^2_{l,\theta}I(x,y) \tag{3}$$

where

$$\nabla^2_{l,\theta}I(i,j)=|2I_{l,\theta}(i,j)-I_{l,\theta}(i-\text{step},j)-I_{l,\theta}(i+\text{step},j)| \\ +|2I_{l,\theta}(i,j)-I_{l,\theta}(i,j-\text{step})-I_{l,\theta}(i,j+\text{step})| \tag{4}$$