Biomedical Signal Processing and Control 47 (2019) 63–74

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

MRI and PET/SPECT image fusion at feature level using ant colony based segmentation

Hamid Reza Shahdoosti*, Zahra Tabatabaei

Department of Electrical Engineering, Hamedan University of Technology, Hamedan 65155, Iran

ARTICLE INFO

Article history:
Received 12 April 2018
Received in revised form 21 July 2018
Accepted 14 August 2018

Keywords:
Medical image fusion
Ant colony algorithm
Object-based fusion
Ensemble empirical mode decomposition

ABSTRACT

Extracting salient features from the medical images and combining them by an appropriate algorithm are the key challenges of multimodal image fusion. The commonly used coefficient-wise fusion may also inject noise into the merged images. To tackle the problem, this paper proposes a new method of multimodal image fusion which makes use of a segmentation map given by the ant colony algorithm. Firstly, the proposed method applies the maximum selection rule in ensemble empirical mode decomposition (EEMD) domain to obtain a fusion map. Then, the proposed approach exploits the color information of the pseudo-color image (PET or SPECT) to find spatial regions of pixels belonging to the same object. This step gives the segmentation map. Finally, the proposed method uses the majority voting process to combine the results of the fusion map and the segmentation map. In fact, the majority voting process determines the winner in each region and scale. The EEMD transform is used to decompose images because it is an adaptive and fully data-driven multiscale transform, and the ant colony algorithm is used for segmentation because it can yield a near optimal segmentation solution. Experimental fusion results are presented on three medical image datasets. It is shown in experiments that the proposed scheme improves the fusion results and provides images with more spatial and color information, when compared to state-of-the-art methods.

1. Introduction

The main aim of multimodal medical image fusion can be generally defined as merging useful information contained in a set of input medical images into a single output image without introducing redundancy, information loss, and distortion [1]. In practice, it is impossible to fully achieve this goal, and thus partially satisfying the mentioned requirements can be considered as a suitable target.

Medical images provide structural or functional information [2]. MRI (magnetic resonance imaging) and CT (computed tomography) images represent structural and anatomical information with high- spatial resolution. SPECT (single-photon emission computed tomography) and PET (positron emission tomography) images represent functional information with low spatial resolution. By the fusion of MRI and pseudo-color (PET or SPECT) images, a more informative image is obtained containing not only functional but also structural information. This image is very useful for noninvasive diagnosis of diseases [2].

Up to now, significant progress on multimodal image fusion approaches has been made, and a wide variety of fusion techniques have been proposed. The three broad levels at which fusion of medical images is performed are decision level, feature level, and pixel level among which feature level fusion has attracted considerable research attention [3–6]. Feature level fusion firstly extracts salient region-dependant features from the input images, and then combines the features to yield a final fused image. A review of medical image fusion methods is given in Section 2.

The EMD (empirical mode decomposition) transform is a completely data-driven and adaptive one without pre-determined basis functions. This transform has been designed for multiresolution decomposition and time-frequency analysis of natural signals [7]. The EMD transform decomposes a signal into a set of intrinsic oscillatory modes called IMFs (intrinsic mode functions). To overcome some limitations of the EMD transform such as mode mixing which is a consequence of signal intermittency, the EEMD transform has been proposed [8]. Due to these advantages, the EEMD transform is used in our fusion framework.

Many research studies have confirmed that region-based image fusion is an effective strategy with high performance in terms of preserving spatial and color information [3–6]. Region-based image fusion methods divide the source images into a set of regions by

* Corresponding author.
E-mail address: h.doosti@hut.ac.ir (H.R. Shahdoosti).

https://doi.org/10.1016/j.bspc.2018.08.017
1746-8094/

means of a suitable clustering algorithm and then fuse them region by region. Among the hundreds of clustering techniques, the ant-colony-based clustering is known as one of the most powerful grouping methods [9]. This algorithm is a nature-inspired optimization one based on the natural collective behavior of ant colonies. Our framework makes use of the ant-colony-based clustering algorithm owing to the fact that this algorithm can yield a near optimal clustering solution [9].

In this paper, we propose to employ a region-based framework for medical image fusion. To the best of the authors' knowledge, it is the first time that the region-based strategy is adopted for the fusion of medical images. The region-based image fusion can reduce sensitivity to noise which is a serious problem in pixel-based image fusion methods. In addition, it is the first time that the ant-colony-based clustering algorithm is used to determine different image regions with the aim of image fusion. In this paper, the image fusion problem is treated as an object based classification task such that the proposed framework exploits the majority voting to determine salient coefficients in each region and scale.

The remainder of this paper proceeds as follows. In Section 2, we review the prior literature in the field of multimodal medical image fusion. Preliminaries such as the two dimensional EEMD transform and ant-colony-based clustering are reviewed in Section 3 (subsections 3.1 and 3.2). Subsequently, Section 3 presents the proposed medical image fusion framework in detail (subsection 3.3). Experimental results using both quantitative and qualitative criteria are presented and discussed in Section 4. Finally, Section 5 comprises the conclusion of this study.

2. Prior literature

Owing to its important role in clinical applications, significant advances have been made in the area of multimodal medical image fusion during the last decades. From the viewpoint of the domain in which the algorithms operate, image fusion techniques can be divided into three general groups: 1) component substitution (CS) based methods, 2) multi-resolution (MR) based methods, and 3) variational methods. Traditional CS-based image fusion methods are intensity-hue-saturation (IHS) [10] and principal component analysis (PCA) [11,12], and traditional MR-based ones are Laplacian pyramid algorithm [13] and wavelet transform [14]. Moreover, methods proposed in Refs [15]. and [16] can be considered as the traditional methods of the variational family. Because the proposed method belongs to the second group, this section focuses on MR-based methods. In addition, because the proposed method divides the source images into different regions before the fusion process, the methods that fuse the images region by region are also reviewed in the following paragraphs.

MR-based fusion approaches such as pyramid transformations, have been used extensively during three recent decades. Most of pyramid-based image fusion methods have been developed around the idea of the Gaussian pyramid transform. The most well-known pyramid transformations are the gradient pyramid [17], Laplacian pyramid [18], FSD pyramid [19], morphological pyramid [20], contrast pyramid [21], and ratio pyramid [22]. The discrete wavelet transform (DWT) was proposed by Mallat which utilizes a simple hierarchical framework for interpreting the image information [23]. Ridgelet [24], curvelet [25], ripplet [26], contourlet [27], shearlet [28], EMD [7], and EEMD [8] transforms are advanced versions of multiscale transforms by which more sparse coefficients are obtained. Among these transforms, the EEMD has adaptive basis functions without sensitivity to small perturbation of data [8].

After decomposing the images into sparse coefficients by an appropriate transform, one should use a fusion rule to select the best coefficients containing salient features. Maximum selection is one of the most simple yet efficient rules used in many references for the selection of informative coefficients in the transform domain [29]. In order to capture most relevant information from source images, Ref [30]. introduced a new fusion rule based on directive contrast for the fusion of low and high frequency components of contourlet subbands. Some methods exploited pulse coupled neural network (PCNN) in order to obtain fused coefficients in the transform domain [31,32]. In the method proposed by Xu [1], the local energy of coefficients was used as a measure of saliency in the coarse layer, and the local contrast between detailed layers and the coarse layer was utilized to select coefficients in detailed layers. To transfer more information from the source images into the synthesized image, a fusion rule was proposed by Wang et al. [28] which fully takes into account inter-scale and intra-scale dependencies between shearlet coefficients. An MRI and PET/SPECT image fusion method was introduced in [33] using a local Laplacian filtering-based technique established through a novel multi-scale system architecture. The local energy maximum scheme is used to obtain fused approximate images and an information of interest-based scheme is used to generate the fused residual images. A new fusion approach was presented by Shen et al. [34] which considers the exposure quality measurement between different exposure images using the local weight, the global weight and the just noticeable distortion (JND)-based saliency weight. In addition, a new hybrid exposure weight is used in this method which is guided by not only a single image's exposure level but also the relative exposure level between various exposure images. For fusion of anatomical and functional images, a novel method was proposed in [35] making use of the combination of parallel saliency features in a multi-scale domain. Firstly, source images are decomposed into a series of smooth layers and detail layers at different scales. Then, the saliency maps are obtained by extracting the parallel saliency features of both sharp edge and color detail. For this purpose, the Canny edge detection operator and the context-aware operator are used. A novel patch-based match and fusion algorithm was introduced in [36] which takes into account moving scene in a multiple exposure image sequence by using optimization. In order to find and match the corresponding patches in various exposure images, a uniform iterative method is used in this fusion method. A multi-scale fusion approach was presented in [37] which can synthesize any number of source images through MEMD (multivariate empirical mode decomposition) transform. The saliency map in this method is obtained by weighted averaging of the absolute values of the pixel intensities within different windows. An MEMD algorithm was introduced in [38] for the pansharpening application. The variance measure is used in this method to quantify the degree of local details of source images that need to be injected into the final fused image. A bivariate EMD-based image fusion approach was proposed by Rehman et al. [39], by which one can combine both gray-level and RGB based color images.

To take advantage of both CS and MR based methods, some hybrid methods were proposed. Because, the IHS algorithm and the retina-inspired model (RIM) fusion technique can preserve more spatial feature and more functional information content, respectively, the method proposed in [2] has combined IHS and RIM approaches (the RIM technique is an MR-based fusion method [2]). To further improve the quality of fused images, the joint use of PCA and DWT was proposed by Naidu and Raol [40]. In the method introduced in Refs [41]. and [42] the contourlet transform and ridgelet transform are used instead of the DWT transform, respectively.

The major problem of choosing informative coefficients based on a single coefficient value is that this strategy is very sensitive to noise. Although the region-based fusion has been rarely used in the field of medical image fusion, this strategy can address the problem. Region-based fusion techniques can be considered as a subset of the feature level image fusion methods [43]. Here, some