Pattern Recognition 79 (2018) 130–146

Contents lists available at ScienceDirect

Pattern Recognition

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

Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning

Huafeng Li [a], Xiaoge He [a], Dapeng Tao [b,c,*], Yuanyan Tang [c], Ruxin Wang [d]

[a] Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, Yunnan 650500, PR China
[b] School of Information Science and Engineering, Yunnan University, Kunming, Yunnan 650500, PR China
[c] Faculty of Science and Technology, University of Macau, Macau 999078, PR China
[d] Yunnan Union Visual Innovation Technology, Kunming, Yunnan 650500, PR China

ARTICLE INFO

Article history:
Received 25 April 2017
Revised 26 November 2017
Accepted 2 February 2018

Keywords:
Medical image fusion
Denoising
Dictionary learning
Low-rank decomposition
Sparse representation

ABSTRACT

Medical image fusion is important in image-guided medical diagnostics, treatment, and other computer vision tasks. However, most current approaches assume that the source images are noise-free, which is not usually the case in practice. The performance of traditional fusion methods decreases significantly when images are corrupted with noise. It is therefore necessary to develop a fusion method that accurately preserves detailed information even when images are corrupted. However, suppressing noise and enhancing textural details are difficult to achieve simultaneously. In this paper, we develop a novel medical image fusion, denoising, and enhancement method based on low-rank sparse component decomposition and dictionary learning. Specifically, to improve the discriminative ability of the learned dictionaries, we incorporate low-rank and sparse regularization terms into the dictionary learning model. Furthermore, in the image decomposition model, we impose a weighted nuclear norm and sparse constraint on the sparse component to remove noise and preserve textural details. Finally, the fused result is constructed by combining the fused low-rank and sparse components of the source images. Experimental results demonstrate that the proposed method consistently outperforms existing state-of-the-art methods in terms of both visual and quantitative evaluations.

1. Introduction

Medical imaging is a fundamental and powerful tool in modern medical treatment and diagnosis, with a broad range of modalities now available including computed tomography (CT), positron emission tomography (PET), and magnetic resonance imaging (MRI). However, different medical imaging modalities - each with their own purposes, strengths, and limitations - convey different information about the observed object, and a single medical imaging modality often cannot provide sufficient information for its intended purpose. Therefore, integrating complementary information from different modalities is a valuable way to obtain more informative results than provided by each primary image alone.

Medical image fusion has emerged as an effective information combining technique, with many approaches proposed over recent years [1–3]. The most popular method is based on multiscale analysis (MTA), and commonly used MTA methods include the discrete wavelet transform (DWT) [4,5], shearlet transform (ST) [6], curvelet transform (CVT) [7,8], contourlet transform [9], and non-subsampled contourlet transform (NSCT) [10]. Generally, these methods first decompose the source image into different subbands at different scales and then fuse these subbands according to fusion rules. Finally, the fused image is constructed using the inverse multiscale transform method. However, the fusion rules used in different subbands do not always successfully identify clear pixels or coefficients, resulting in the loss of detailed information and degradation of image quality.

Sparse representation (SR) theory and dictionary learning (DL) have also been used for image fusion [1,3,11–14]. SR-based medical image fusion first finds the sparsest representation of source images over a given dictionary, before integrating the representation coefficients according to a fusion rule. Finally, the fused image is constructed by combining the given dictionary and the fused sparse coefficients. Although the SR-based method has been successfully applied to medical image fusion, only one dictionary is usually employed to represent different morphological parts of the source images, resulting in suboptimal representation of inherent information in the source images. To this end, Jiang and Wang [15] proposed a strategy to separate the source images into cartoon

* Corresponding author at: Chenggong Daxuecheng, Kunming, PR China.
E-mail address: dapeng.tao@gmail.com (D. Tao).

https://doi.org/10.1016/j.patcog.2018.02.005

and textural components for image fusion using DCT and curvelet dictionaries.

In spite of the success of medical image fusion, the methods mentioned above tend to have poor fusion performance in the presence of noise, since the assumption is that the source images are noise-free [16]. However, there is always some acquisition or transmission noise in medical images in practice. To overcome this problem, one commonly used strategy is to first denoise the source images before fusing them. By doing so, the generated artifacts from image denoising may be propagated and magnified during fusion. Moreover, such a sequential approach can cause further information loss. Hence, a method that simultaneously conducts image fusion and denoising would be valuable. Some simultaneous approaches have been proposed [17,18], but suppressing noise and preserving textural details remains challenging.

Considering that an image can be modeled as a superposition of low-rank and sparse components and the low-rank representation can remove noise from corrupted observations, here we propose a novel medical image fusion, denoising, and enhancement approach based on low-rank sparse component decomposition and dictionary learning. Our method is inspired by pattern classification based on dictionary learning, and the image decomposition problem is treated as a classification. To endow the low-rank and sparse component dictionaries with powerful discriminative capability, we enforce low-rank and sparse constraints on the low-rank and sparse components, respectively. Furthermore, to more effectively perform source image decomposition, we formulate a novel sparse coding model over the learned dictionaries as a joint estimate of the low-rank and sparse components. Specifically, noise is removed from the noisy image by imposing weighted nuclear norm regularization on the separated sparse component.

Moreover, we incorporate a sparse constraint term into the image separation model for the sparse component to enhance textural detail. After decomposing the source images into low-rank and sparse components, we can obtain their representation coefficients over the corresponding dictionaries. During fusion, the coefficients of the same component originating from different imaging modalities are integrated by the "max-absolute" fusion rule. Then, the fused low-rank sparse components are reconstructed by combining the fused coding coefficients and their respective dictionaries. Finally, we construct the final fused image by superposing the two fused components. The contributions of our work and the advantages of our joint framework can be summarized as follows:

(1) We formulate an efficient multiple component dictionary learning model by integrating rank minimization into the low-rank component. To further improve the discriminative ability of the learned dictionaries, a sparse constraint is imposed on the sparse component.
(2) In the image decomposition model, a weighted nuclear norm regularization and nuclear norm constraint term are introduced to significantly retain edge information during noise removal. To enhance the visual effects during fusion, sparse regularization of the sparse component is also exploited.
(3) We propose using an alternating iteration method to produce the coding coefficients. Consequently, our method effectively preserves detailed information during fusion and denoising. Moreover, our approach can conduct fusion, denoising, and enhancement simultaneously. As a result, there is no propagation of artifacts as seen with multi-step approaches.

The remainder of the paper is organized as follows. In Section 2, we introduce related work including DL and its application in image fusion. In Section 3, we formulate the dictionary learning, sparse representation, and low-rank decomposition of input images. Section 4 presents the dictionary learning and low-rank sparse decomposition model optimization algorithms. Section 5 describes medical image fusion. Experimental results of fusion of multi-modality medical images are presented in Section 6, and we conclude and discuss future research in Section 7.

2. Related work

2.1. Dictionary learning

In SR and low-rank learning, the selection of dictionary $\boldsymbol{D}$ is important and critical because the performance of these learning models is highly dependent on it. Generally, $\boldsymbol{D}$ can either be analytically designed or learned from a training sample. With respect to image fusion, Jiang and Wang [15] chose curvelets as a dictionary to represent the cartoon content due to their outstanding ability to detect smooth curves, edges, and anisotropic structures, and selected a DCT dictionary to express the textural content because it can represent periodic structures such as texture sparsely. However, such analytical dictionaries do not adaptively and sufficiently characterize natural image structures [19]. A dictionary learned from example image patches is a more commonly used approach, the typical algorithm being KSVD [20]. For image restoration, Dong et al. [19] utilized *K*-means clustering to cluster the training patches and then learned a dictionary of PCA bases for each cluster, and, to suppress noise, exploited non-local self-similarity [21] of image patches.

DL has also been applied to pattern classification and recognition. For image classification, Bahrampour et al. [22] developed a multimodal task-driven dictionary-learning algorithm by utilizing the joint sparsity constraint to enforce collaborations among multiple homogeneous/heterogeneous information sources. Akhtar et al. [23] presented a Bayesian approach for learning discriminative dictionaries to build associations between the dictionary atoms and class labels. For improved discriminability and more robust classification, Quan et al. [24] proposed a discriminative sparse coding method by simultaneously learning a dictionary and training an ensemble classifier. For robust image classification, Rong et al. [25] presented low-rank double dictionary learning, which learned a discriminative dictionary from corrupted data.

2.2. DL and SR in image fusion

For image fusion, Kim et al. [11] proposed a sub-dictionary constructing method for each cluster. Once all cluster sub-dictionaries were learned, a complete dictionary was constructed by merging the training results. In [3], Li et al. presented a medical fusion method using group-sparse representation and DL. Based on the density peaks clustering algorithm, Zhu et al. [1] developed a novel compact DL approach for multi-modal medical image fusion. To capture the intrinsic characteristics of images and preserve the hierarchical structure of stationary wavelets, Yin [26] proposed a joint DL strategy for all stationary wavelet subbands. However, this approach was MTA-based, so was not robust to noise or changes in fused image coefficients. Furthermore, the above DL-based methods represented different components of source images through one dictionary. Images often contain structures with different spatial morphologies [15]. If one wants to reconstruct a better quality image, larger dictionaries and more training samples are needed, which affects algorithm efficiency [27].