Pattern Recognition Letters 100 (2017) 22–28

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Classification of nematode image stacks by an information fusion based multilinear approach

Min Liu[a,*], Xueping Wang[a], Hongzhong Zhang[b]

[a] College of Electrical and Information Engineering, Hunan University, Hunan 410082, China
[b] Department of Industrial Engineering & Operations Research, Columbia University, 1255 Amsterdam Ave, New York, NY 10027, United States

ARTICLE INFO

Article history:
Received 28 January 2017
Available online 14 September 2017

ABSTRACT

In this letter, we present to use an information fusion based multilinear analysis approach to classify multi-focal image stacks. First, image fusion techniques such as the nonsubsampled contourlet transform sparse representation (NSCTSR) are used to combine relevant information of multi-focal images within a given image stack into a single image, which is more informative and complete than any single image in the given image stack. Second, multi-focal images within a stack are fused along 3 orthogonal directions, and multiple features extracted from the fused images along different directions are combined by using canonical correlation analysis (CCA). Finally, because multi-focal image stacks represent the effect of different factors - texture, shape, different instances within the same class and different classes of the objects, we embed the information fusion methods within a multilinear analysis (MA) framework to propose an information fusion based multilinear classifier. The experimental results demonstrated that the information fusion based multilinear classifier can reach a higher classification rate (96.6%) than the previous multilinear based approach (86.4%), even we only use the texture feature instead of the combination of texture and shape features as in the previous work.

1. Introduction

In biological or medical field, digital multi-focal images (DMI) are very important for documentation and communication of specimen data, because the morphological information for a transparent specimen can be captured in the form of a stack of high-quality images, representing individual focal planes through the specimen's body [1]. Four multi-focal image stacks taken from a differential interference contrast microscope are shown in Fig. 1. Each stack contains multiple focal planes taken from the top to the bottom of the specimen, with only a few frames of each shown. Given such image stacks containing so many multi-focal images, how do we efficiently extract effective features from all layers to classify the image stacks is still a problem.

Most of the work on image feature extraction and classification are done through the 2D image processing methods [2–9,33]. There is very little work about the topic of 3D image stack classification. Existing 3D feature extraction methods like 3D scale-invariant feature transform (SIFT) and 3D histogram of oriented gradient (HOG) [10,11] are not suited for this purpose because images in multiple focal planes have different characteristics [16].

Previously, a projection based multilinear classifier is presented to classify the nematode multi-focal images [16], where the 3D X-Ray Transform is used for the projection of the multi-focal image stacks. However, in the projection based multilinear classifier, the features extracted from the projection images along oblique directions are not very reliable for classification, especially when the precision along the *z*-direction is not high enough. It poses limitation to the multi-direction projection based multilinear classifier.

In this letter, we propose to use the information fusion based multilinear classification approach to classify multi-focal image stacks. On one hand, the image fusion methods can combine relevant information from multiple images of a given image stack into a single image, and the resultant fused image will be more informative and complete than any individual focal plane images. On the other hand, for the classification of multi-focal image stacks, there are multiple factors that need to be analyzed–texture, shape, different instances within the same class and different classes of specimens [16]. For this purpose, we embed the information fusion methods within a multilinear framework, as illustrated in Fig. 4.

In fact, the multilinear image analysis method plays an important role in the recognition of natural face images formed by different factors, such as expression, illumination conditions, pose and so on [18–21]. In multilinear analysis, the facial image ensemble is represented as a high dimensional tensor. By a high-order gen-

* Corresponding author.
E-mail address: liu_min@hnu.edu.cn (M. Liu).

http://dx.doi.org/10.1016/j.patrec.2017.09.024
0167-8655/

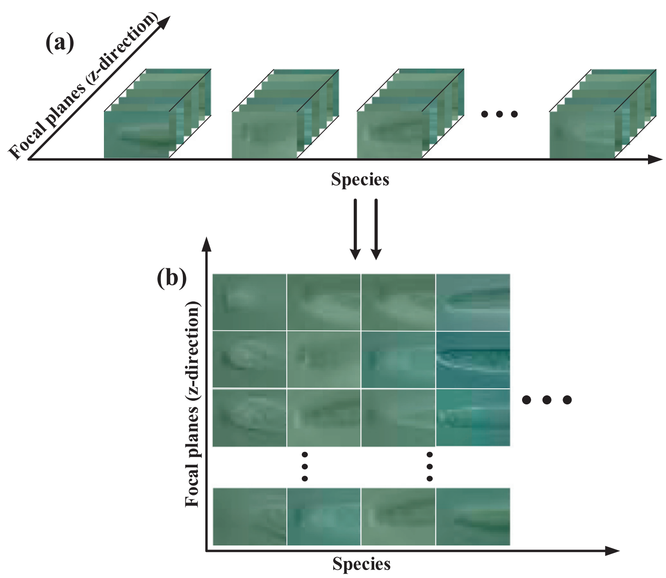

Fig. 1. A few samples of multi-focal image stacks taken from different nematode species (horizontal direction) are shown in (a) and (b). Each stack contains images of multiple focal planes taken from the top to the bottom of the specimen, with only a few frames of each shown (vertical direction).

eralization of principal component analysis (PCA) and high-order singular value decomposition (HOSVD) [12,13,17,18], the collection of facial images is represented by separating the different modes underlying the formation of facial images. This idea is applied to the classification of multi-focal image stacks, since they also represent the combination of various factors [16], such as texture, shape, different instances within the same class, and different classes of specimens and so on. The multilinear analysis allows us to model the inter-relationships between the different variables, while at the same time allowing independent analysis along each dimension.

Besides, multi-focal images within a given stack are fused along 3 orthogonal directions, and multiple features extracted from the fused images along different directions are combined using canonical correlation analysis (CCA) [14,15], and then the combined features are provided for the multilinear analysis framework.

The proposed information fusion based multilinear classifier is applied in the classification for multi-focal image stacks of nematodes – a species which is very difficult to classify since they are one of the most numerous animals on earth [1]. Because the texture information is more robust than the shape information in most of the nematode images, in this letter we mostly show how to derive features for multi-focal image data sources that take into account the texture information only.

2. Methodology

2.1. Multi-scale transform sparse representation (MSTSR) image fusion

In multi-focal image stacks, high-quality images usually contain complementary information which could be integrated. Through image fusion process extended or enhanced information content can be obtained in the composite image, which could be extremely useful for image classification application. In this letter, we employ a multi-scale transform sparse representation image fusion method to fuse the multi-focal image stacks for image classification, which could take the complementary advantages of multi-scale transform (MST) and sparse representation (SR).

The MST-based image fusion methods assume that the underlying salient information of the source images can be extracted from the decomposed coefficients. Its basic idea is illustrated in Fig 2. Given two source images I_A and I_B, it consists of the following three steps in general [22]. First, the source images are decomposed into a multi-scale transform domain with low-pass bands $\{L_A,\ L_B\}$ and high-pass MST bands $\{H_A,\ H_B\}$; second, the transformed coefficients $\{L_A,\ L_B,\ H_A,\ H_B\}$ are merged with a given fusion rule; finally, the fused image $I_F$is reconstructed by performing the corresponding inverse transform over the merged coefficients.

For sparse representation based image fusion algorithm [23], the sparse coefficients are used as the local salient feature and the coefficients are merged with the "choose-max" fusion rule using al_1- norm. The fused image is then reconstructed using the merged coefficients with a learned dictionary $\mathbf{D}$ [27–30]. The whole image is processed by using the "sliding window" technique, which is a widely-used technique in various SR based image processing applications.

Although both the MST- and SR-based image fusion methods have achieved great success in image fusion, it is worthwhile to notice that both of them have some defects [22]. In fact, an image fusion framework by taking the complementary advantages of MST and SR is presented to overcome the related disadvantages [22].

Here we choose the nonsubsampled contourlet transform (NSCT) for illustration. The diagram of the nonsubsampled contourlet transform sparse representation (NSCTSR) image fusion method for two source images I_A and I_B is shown in Fig. 2.

Specifically, the high-pass NSCT bands H_A and H_B are fused using the conventional "max absolute" rule shown in Eq. (1),

$$H_F = \max\{H_A, H_B\} \tag{1}$$

H_F is the fused high-pass bands.

For the low-pass NSCT bands L_A and L_B, they are first divided into patches p_A^i and p_B^i by a sliding window technique. Let us assume$\hat{\mathbf{v}}_A^i$ and $\hat{\mathbf{v}}_B^i$ are the normalized result of column vectors $\mathbf{v}_A^i$ and $\mathbf{v}_B^i$from p_A^i and p_B^i, the sparse coefficient vectors γ_A^i and γ_B^i are then computed using the orthogonal matching pursuit (OMP) algorithm [22] as below,

$$\begin{cases} \gamma_A^i = \arg\min\limits_{\alpha} \|\gamma\|_0 s.t. \|\hat{\mathbf{v}}_A^i - \mathbf{D}\gamma\|_2 < \varepsilon \\ \gamma_B^i = \arg\min\limits_{\alpha} \|\gamma\|_0 s.t. \|\hat{\mathbf{v}}_B^i - \mathbf{D}\gamma\|_2 < \varepsilon \end{cases}, \tag{2}$$

Thenγ_A^i and γ_B^i are merged with a SR-based fusion approach (max l_1- norm), as shown in Eq. (3),

$$\gamma_F^i = \begin{cases} \gamma_A^i \ if \ \|\gamma_A^i\|_1 > \|\gamma_B^i\|_1 \\ \gamma_B^i \quad \text{otherwise} \end{cases}, \tag{3}$$

where $\gamma_F^i$is the fused sparse vector. The fused result of $\mathbf{v}_A^i$ and $\mathbf{v}_B^i$is calculated by

$$\mathbf{v}_F^i = \mathbf{D}\gamma_F^i + \bar{v}_F^i \cdot \mathbf{1} \tag{4}$$

where the merged mean value $\bar{v}_F^i$ is obtained by

$$\bar{v}_F^i = \begin{cases} \bar{v}_A^i \ if \ \gamma_F^i = \gamma_A^i \\ \bar{v}_B^i \quad \text{otherwise} \end{cases} \tag{5}$$

The low-pass fused result $L_F$is then generated by reshaping all $\mathbf{v}_F^i$ s and plugged into the original positions.

Once the fused high-pass bands and low-pass bands are calculated, the fused image I_F is finally obtained by performing the inverse NSCT on the merged coefficients $\{L_F,\ H_F\}$.

The fusion results of the nematode multi-focal image stacks shown in Fig. 1 along the z-direction (as the vertical direction illustrated in Fig. 1) by some typical image fusion methods such as NSCT, SR, dual-tree complex wavelet transform sparse representation (DTCWTSR), laplacian pyramid sparse representation (LPSR), discrete wavelet transform sparse representation (DWTSR), and NSCTSR are shown in Fig. 3 [22,24,25]. In fact, the image fusion