

Study on Identification Method of Foreign Fibers of Seed Cotton in Hyper-spectral Images Based on Minimum Noise Fraction^{*}

Laiqi Xu¹, Xinhua Wei^{2,**}, Xinyun Zhou², Dazhi Yu², and Jinmin Zhang²

¹ School of Electrical & Information Engineering, Jiangsu University, Zhenjiang, 212013, China

² Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education & Jiangsu Province, Jiangsu University, Zhenjiang 212013, China
wei_xh@126.com

Abstract. In order to improve the recognition accuracy of seed cotton foreign fibers, a study on identification method in hyper-spectral images based on Minimum Noise Fraction (MNF) was proposed, which was applied to feature extraction to reduce the dimension of hyper-spectral images. This method reduced the numbers of hyper-spectral data, lessened the images noise to the minimum, but also decreased the computational requirements for subsequent processing. The white foreign fibers and cotton which were in small discrimination were selected in this paper as the research object. The hyper-spectral images were displayed in software ENVI with 256 bands in the wavelength range of 871.60nm-1766.32nm. Afterwards, the images would be processed with the iteration threshold segmentation method, inflation and corrosion. Meanwhile, the correlation of template images and destination images were calculated to find the spectral peaks so that to make template matching to eliminate the images of the cotton seeds. Results of experiments show that the above methods is suitable for identifying foreign fibers of seed cotton which achieved 84.09% rate of recognition.

Keywords: seed cotton, foreign fibers, MNF, hyper-spectral data, dimensionality reduction, feature extraction, template matching.

1 Introduction

The provision of GB1103-1999 is that foreign fibers refer to the non-cotton fibers and colored fibers which have seriously affected the quality of cotton, such as chemical

^{*} Foundation items : Supported by a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions(Jiangsu financial education (2011) No. 8)), the key laboratory of agricultural equipment intelligent high technology research in Jiangsu (BM2009703), and the Program for New Century Excellent Talents in University(NCET-09-0731).

^{**} Corresponding author.

fiber, hair, silk, linen, dyed rope, plastic film, plastic rope (rope, cloth) and so on[1]. The harm caused by the foreign fibers in the textile business is huge, its performance is to reduce the production efficiency of cotton mixed with foreign fibers first. Meanwhile, it is easy to be labeled as numerous small fiber blemish, not only is difficult to remove in textile processing, but will innocently influence the quality of cloth dyeing processes that pulled off shorter and finer fiber. The formation of a large number of small fibrous blemish pinning can easily result in the efficiency weaving [2]. Therefore, in order to realize the automatic sorting of foreign fibers in cotton, a variety of techniques are tried. For example, Liu Huanjun used spectral reflectance to classify of soils[3], where machine vision technology is the primary means of on-line detection of foreign fibers in cotton.

Using the spectral analysis technology distinction sample type is a fast and accurate method. For example, Chenyang created that the characteristic spectra of 3 kinds of gases and the weights was classified by similarity analysis [4]. These methods all directly realized the variety distinction using the test specimen spectrum characteristic peak. However, machine-based visual of the cotton foreign fibers used in indirect detection is the spectral characteristics of the material. Thus, Li Bidan in Tsinghai University examined the colorless plastic, the jute, the white hair silk by using the infrared absorption characteristic of foreign fiber [5]. Jia Dongyao and Ding Tianhuai produced the differences of foreign fibers of composition, physical and chemical parameters [6]. Church J.S. et al. used the near-infrared wave Bands within 1000nm-1700nm to realize in the wool polymer impurity examination [8]. Hyper-spectral image technologies for the detection of foreign fibers in cotton are not much visible in the literature at home and abroad. By analyzing the foreign fibers in the near-infrared hyper-spectral transformation algorithm to choose the best band of foreign fibers in cotton, has not been found in Chinese literature.

This paper takes advantage of the high-resolution digital camera system to get the high spectral images between seed cotton and 5 kinds of white foreign fibers in the wavelength range of 871.60nm-1766.32nm. Hyper-spectral imaging overcomes the limitation of the traditional single-band, as well as the band range limitations which have a relatively narrow band interval. And more number of characteristics of the band are able to get the spectral space features continuous and fine spectral characteristics. Because of the large amount of letter of hyper-spectral, multiband and high redundancy of the feature makes its information processing difficult, and data dimensionality reduction problem has been the application of hyper-spectral information processing problems [9,11]. Yang Wenzhu et al.in China Agricultural University proposed a method for selecting optimal detection bands, based on the extreme distribution of reflectance differences between cotton fiber and foreign fibers[12]. On one hand, data dimension reduction can make the images away from the noise and improve the data quality .On the other hand, removing the worthless band image can reduce the amount of computation and the band number of purposes, as well as the image processing efficiency. To effectively use of hyper-spectral data, dimension reduction is an integral part. The MNF transformation would solve the problem which mentioned.

2 Research Data

In this study, seed cotton and five typical white fibers including white confetti, white hair, polypropylene yarn, and white chemical silk, white plastic were selected. This paper takes advantage of the high-resolution digital camera system to get the high spectral images at the range of 871.60nm-1766.32nm. Each of the band is 4.3nm with 256 bands.

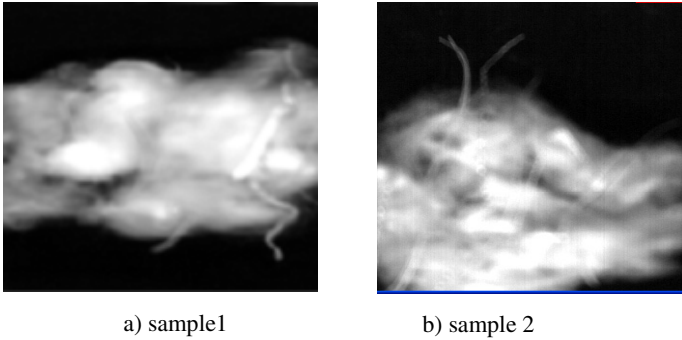


Fig. 1. Foreign fiber samples in seed cotton

3 Research Methods

3.1 The Minimum Noise Fraction (MNF) Transformation

MNF (Minimum Noise Fraction) is one of the common methods of feature extraction in hyper-spectral data [13]. The MNF transformation is essentially two cascaded principal components(PC) transformations[14]. First, the high-pass filter is used in processing the whole image data to get the noise covariance matrix C_N and digitalization.

$$D_N = U^T C_N U \quad (1)$$

In the formula, D_N is the diagonal matrix of the C_N characteristic value which is according to the size descending sequence arrangement. U is the orthogonal matrix which is constituted by the C_N characteristic vector, N is the wave band number. The formula (1) can be further transformed into

$$I = P^T C_N P \quad (2)$$

Where I is the unit matrix, P is the transformation matrix, which $P = U D_N^{-1/2}$. The original image X can be transformed to the new space Y . The data after transforming contained in the noise is with unit variance and without correlation between bands. Standard PC transformation is carried in images. Transformation formula is

$$\hat{C} = P^T C_D P \tag{3}$$

C_D is the covariance matrix of the original image, \hat{C} is the transformed matrix. The above equation can be further transformed into a diagonal matrix:

$$\hat{D} = J^T \hat{C} J \tag{4}$$

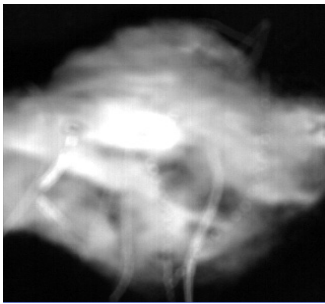
According to the size descending sequence arrangement constitution, \hat{D} is the diagonal matrix of \hat{C} characteristic value, J is the orthogonal matrix which constitutes by the \hat{D} characteristic vector.

Through the above steps, MNF is got to the transformation matrix:

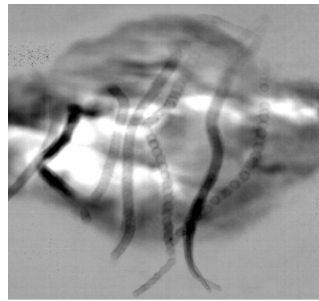
$$T_{MNF} = P J \tag{5}$$

Using the MNF Transformation can remove noise from data by performing a forward transformation and determine which bands contain the coherent images by examining the images and eigenvalues.

As shown below (Fig.2) is the images of foreign fiber before and after by MNF transformation.



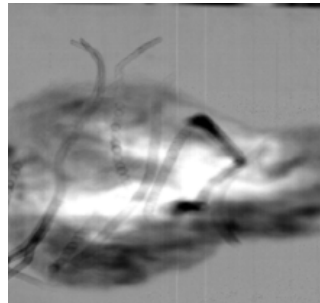
a) The original image



b) Image processed by MNF



c) The original image



d) Image processed by MNF

Fig. 2. Images processed before and after by MNF

3.2 Iteration Threshold Segmentation

For the online visual examination system, it is the key aspects with algorithm rapidity and accuracy. Therefore, choosing a quick division method is necessary under the guarantee division precision. It is more effective for border-based segmentation method in dealing with some relatively high contrast edges images. However, the content of foreign fibers in cotton is extremely small, and became finer after opening. With the low contrast between the target and background, segmentation method based on boundary does not apply to processing the cotton fibers images. Due to the gray of the cotton fiber image varies from the type of foreign fibers, the fixed threshold method can not achieve precise segmentation. In this case, dynamic threshold method is needed. Iterative threshold segmentation method is statistically significant for the best segmentation threshold, which is more suitable for cotton fiber.

The principle of the iterative method is that the gray distribution between foreground and background in the images do not overlap with each other. With this premise, the threshold segmentation is realized in two types of objects based on the idea of approximation. In General, the gray images within the change between 0-255 gray-scale changes is 256 levels. Some steps are given below:

Step1: Find out the maximum and minimum gray value of images to record as Z_{max} and Z_{min} respectively. Thus, make the initial threshold $T_0 = (Z_{max} + Z_{min}) / 2$;

Step2: According to the threshold T_K , to do image segmentation for the foreground and background to get the average gray values Z_o and Z_B •

Step3: Get the new threshold for $T_{K+1} = (Z_o + Z_B) / 2$;

Step4: If $T_K = T_{K+1}$, the income is for threshold; Otherwise, turn to the Step2 to do the iterative calculation.

3.3 Template Matching, Corrosion and Dilation

Template matching is an important part of digital image processing. After the image processing, cotton seed in-depth can have the remarkable images characteristic, in this case, a method that eliminate cotton seed image is needed. Establishing a template of the average size model and carrying on the match with the image is named template match. In this article, the average value of cotton seed template was first found out, and then pattern recognition method was used. Meanwhile, two dimensional Fourier were transformed between the template image and matching image, between which the correlation was calculated .The highest position of the spectrum that match template objects.

After the MNF processing, iterative processing, corrosion and dilation, the foreign fibers were divided with the division effect achieves 84.09%. Now enumerates two pictures to take the demonstration, the demonstration is as follows:

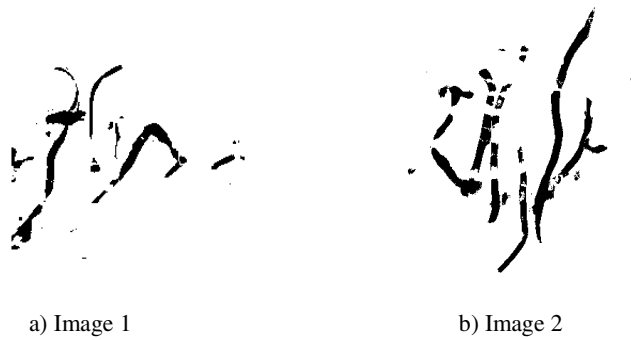


Fig. 3. Final segmentation images

4 Results and Analysis

4.1 The Results of Minimum Noise Fraction (MNF) Transformation

In the figure 4, there are the eigenvalue of the MNF and 10 top features images of the cumulative variance, which is generated by a MNF transformation.

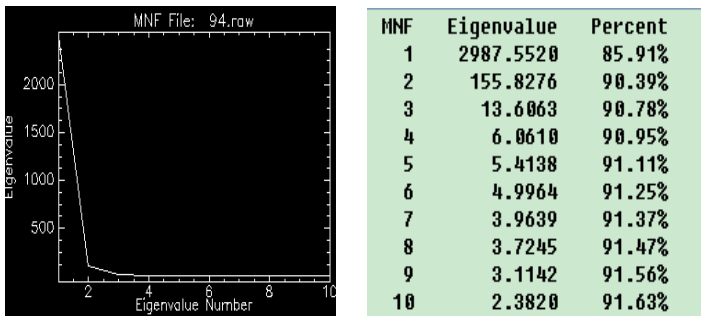


Fig. 4. Eigenvalue of the MNF and 10 top features images of the cumulative variance

It is known that the transformed image of cumulative variance reaches top 10 bands of 87.92%. And most amounts of images are contained in top 10 bands. Typically, after the MNF transformation, the band that contains the data Eigenvalues is greater than 1. When the Eigenvalue is close to 1, it indicates that the images have only noise. As we can see in figure 4, the fifth band characteristics are 1, which contains part of the noise. The first 4 after MNF processing components is shown in Figure 5.

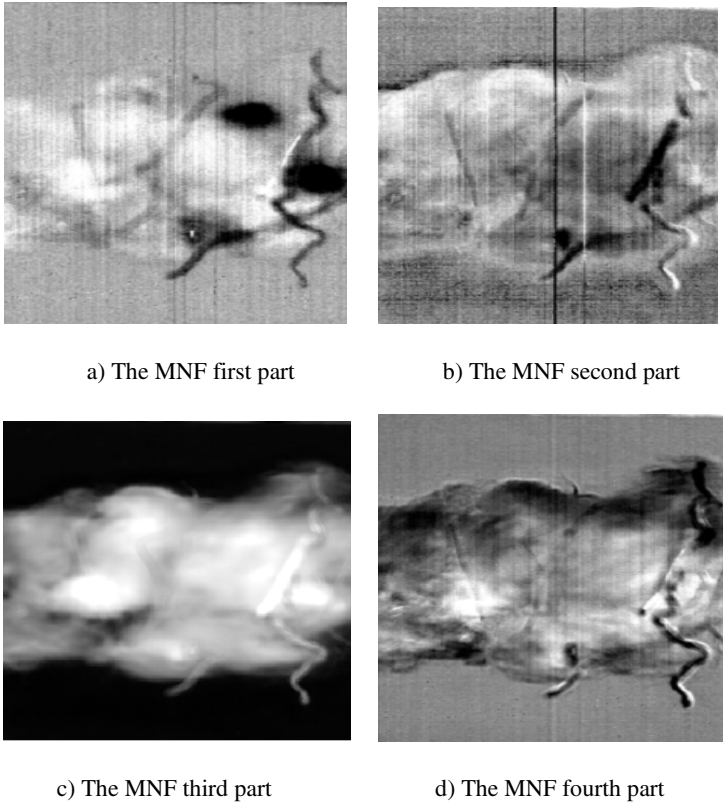


Fig. 5. The Results of MNF

4.2 Results of Iteration Threshold Segmentation

In order to deal with further better in the image processing, appropriate image segmentation is needed. First of all, the iterative method is imported into the MATLAB to figure an iterative process. Combined with histograms, the threshold values used for the calculation of the iterative method are proposed. The method is based on 256 gray-scales of foreign fiber images, and the initial value of the last iteration is selected.

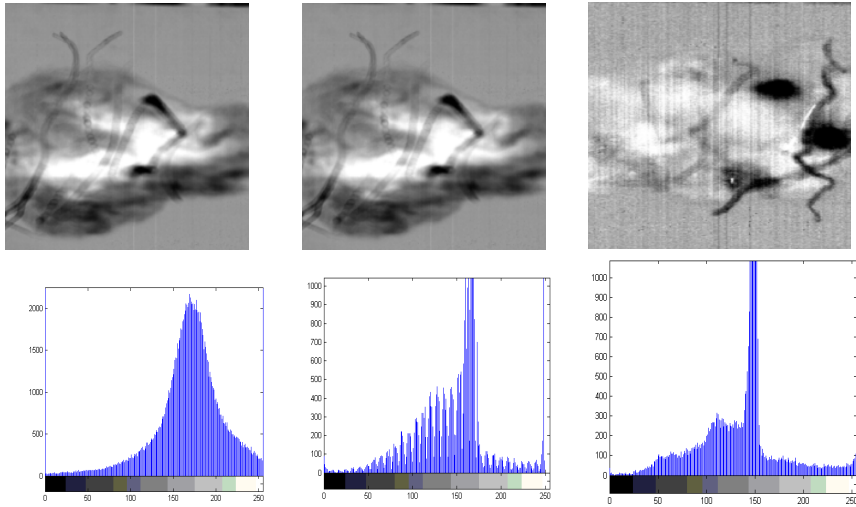


Fig. 6. Gray images 1,2,3 and its histograms

Table 1. The results of optimal threshold

Image number	Image 1	Image 2	Image 3
Optimal threshold	154	139	112

According to the principles of iterative methods, iterative calculation should be done. Finally, the thresholds are as shown in table 1.

4.3 Results of Template Matching

To calculate correlation of template image and destination image, the template matching should be selected. They are listed as follows.

Step1: Convert into a template (396,398) binary image, and then do a matching image to convert a (396,398) binary image.

Step2: To do two-dimensional Fourier transformation of the template image and match the image.

Step3: Calculate correlation of template image and destination image. The method is to match the image rotated 180 degrees, then convolution calculation techniques based on fast Fourier transform evaluated. If the convolution rotates 180 degrees, the convolution and related calculations are equivalent.

Step4: Find one of the three highest spectral peaks from the generated spectrum image spectral peaks. The three highest positions of the spectral peaks are matched with the template object.

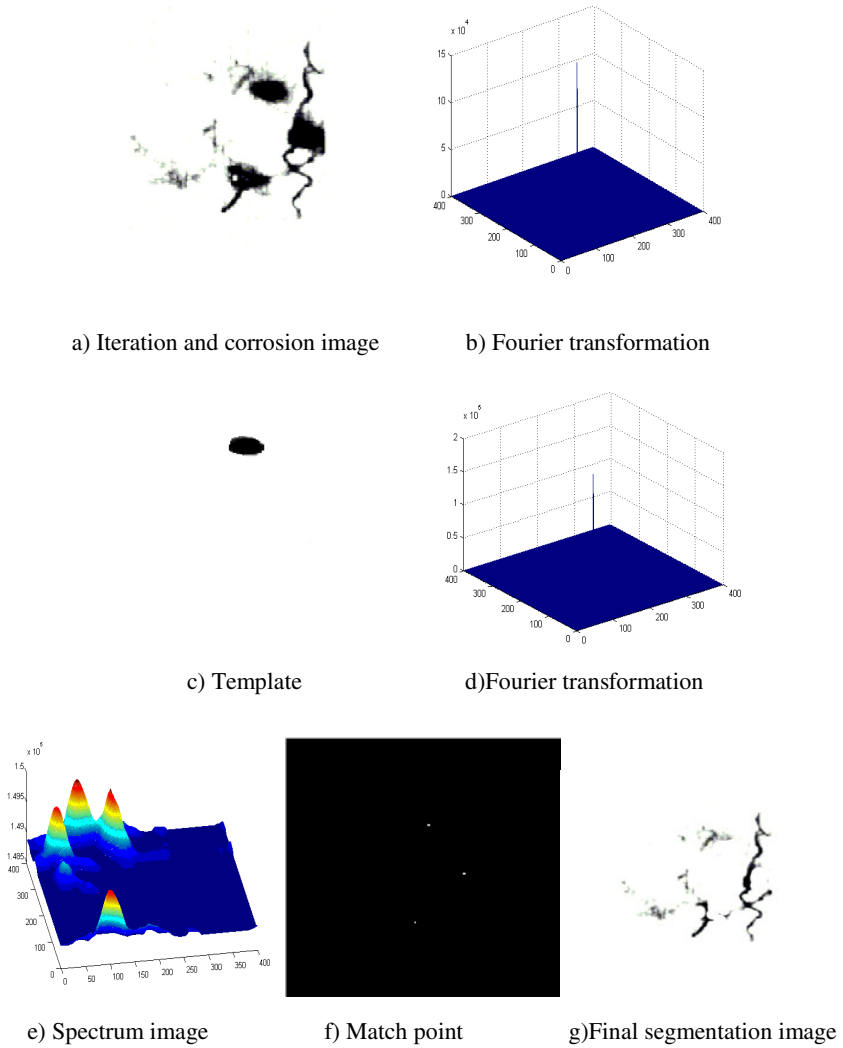


Fig. 7. The result of segmentation process

4.4 Data Statistical Results

With the proposed method, segmentation and dimensionality treatment has been used in the hyper-spectral images of the cotton foreign fiber, in order to show the method, there are 176 sets of data tests, and the experimental results are expressed as follows.

Table 2. Results of testing sets of samples

Name	Sample number	Correct identification number	Correct identification (%)
white confetti	32	30	93.75
white hair	39	29	74.35
polypropylene silk	36	32	88.88
white chemical silk	39	35	89.75
white plastic	30	22	73.33
Sum	176	148	84.09

5 Conclusions

Through theoretical analysis and experimental analysis of this article, we find that Green and other people's classic MNF transformation methods have the good effects in dimensionality reduction, separation of the of noise and feature extraction for similar surface features gathered data. Currently result fully illustrates the advantages of MNF transformation method. However, for the time complexity of the algorithm of hyper-spectral images, it is necessary for us to do further simplifying and further mining algorithm.

References

- [1] Fang, H.: Analysis of application for the MQZY seed cotton foreign fibers. *Cotton Processing in China* (4), 12–14 (2010)
- [2] Li, M.: Features and application of online foreign fiber detecting and clearing device. *Shanghai Textile Science&Technology* 34(1), 15–18 (2006)
- [3] Liu, H., Zhang, B., Zhang, Y., et al.: Soil taxonomy on the basis of reflectance spectral characteristics. *Spectroscopy and Spectral Analysis* 28(3), 624–628 (2008)
- [4] Chen, Y., Jia, L., Zhang, T., Guo, P., Wang, X., Chang, S.: A classification method for nonlinear fluorescent spectra based on edges matching. *Acta Physica Sinica* 59(1), 271–279 (2010)
- [5] Li, B., Ding, T., Jia, D.: Design of a sophisticated foreign fiber separator. *Transactions of the Chinese Society for Agricultural Machinery* 37(1), 107–110 (2006)
- [6] Jia, D., Ding, T.: The use of fiber infrared absorption characteristics of lint impurities detection method. *Infrared and Millimeter Waves* 24(2), 147–150 (2005)
- [7] Xia, X., Ji, J., Chen, J., Liao, Q., Yang, Z.: Analysis of soil physical and chemical properties by reflectance spectroscopy. *GEO-edge* (4), 354–362 (2009)
- [8] Church, J.S., O'Neil, J.A.: The detection of polymeric contaminants in loose scoured wool. *Vibrational Spectroscopy* 19(2), 285–293 (1999)
- [9] Tong, Q., Zhang, B., Zheng, L.: *Hyper-spectral remote sensing principles techniques and applications*. Higher Education Press, Beijing (2006)

- [10] Hughes, G.F.: On the Mean Accuracy of Statistical Pattern Recognition. *IEEE Trans. Info. Theory* 14(1), 55–63 (2008)
- [11] Du, P., Chen, Y.: *Advances in remote sensing science*, pp. 369–370. Xuzhou China mine University Press (2007)
- [12] Yang, W., Li, D., Wei, X., et al.: Selection of optimal band for detecting foreign fibers in lint cotton using spectroscopic analysis. *Transactions of the CSAE* 25(10), 186–192 (2009)
- [13] Liu, P., Lin, H., Sun, H., Yan, E.: Dimensionality reduction method of Hyperion Eo-1 data. *Journal of Central South University of Forestry & Technology* 31(11), 34–38 (2011)
- [14] Green, A., Berman, M., Switzer, P., Craig, M.D.: A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing* 26(1), 65–74 (1988)