Image classification of cashmere and wool fiber based on LC-KSVD

Chunhong Sun
Shanghai University, Shanghai, 200444, China
Email: sun_flower@shu.edu.cn

Abstract. The identification of cashmere and wool fibers has always been a key issue in the textile industry. The classification method of cashmere wool based on microscope image processing technology mainly studies fiber structure characteristics, and the steps are complicated. Affected by changes in breeding methods, breed improvements and climate change, the output of cashmere has decreased, and cashmere fibers have more and more variations, which makes the classification of cashmere wool more difficult. In order to accurately identify cashmere and wool fibers, this paper proposes a multi-feature fusion cashmere wool classification algorithm based on the LC-KSVD algorithm. According to the characteristics of the microscope fiber image, the key information and texture features of the fiber image are extracted, and then the multiple types of texture features are merged into Multi-dimensional feature matrix vector, and finally realize the classification of cashmere wool through the LC-KSVD algorithm. Experiments show that compared with other classification algorithms, the classification accuracy of this method can reach up to 90%, which can be used for subsequent cashmere wool fiber classification and identification.

1. Introduction

In the textile industry, cashmere wool fiber has a wide range of applications, has high practical value, and is a very important animal fiber. When studying individual cashmere wool fibers based on the characteristics of cashmere wool fibers, due to changes in breeding methods, breed improvement and climate change, the appearance and morphological characteristics of cashmere wool fibers such as scale information vary: the difference between fiber types is small, the difference within the class is large, and the difference in scale information between some wool fibers and cashmere fibers becomes smaller, which makes the discrimination more difficult. Currently commonly used fiber identification technologies include: microscope technology [1], chemical analysis technology [2], DNA analysis technology [3] and microscope-based image processing technology, etc., mainly research fiber structure characteristics and classification and identification applications. Aspects of the problem. There are various methods for identifying cashmere and wool fibers, but different identification methods for research in various fields have their own limitations. There is no identification method that satisfies all research fields. The use of optical microscope to identify cashmere and wool fibers is the earliest and most commonly used detection method, which is convenient to use and low in cost. However, when using microscopy to identify cashmere wool, inspectors need to be proficient in distinguishing similar factors such as the various appearance and morphological characteristics of the two fibers in order to obtain accurate identification results. Therefore, it is necessary for the inspectors to have a high level of inspection and rich experience. How to improve the detection technology and level, and provide the
inspectors with convenient and effective detection methods, is also a relatively realistic and urgent problem at present.

In this paper, by studying the similarity factors of cashmere and wool fibers, using image processing technology to study the optical microscope images of the two fibers, a multi-feature fusion cashmere wool classification algorithm based on the LC-KSVD algorithm [4] is proposed. This method does not require tedious steps to get better results. The algorithm steps are as follows: First, perform the image preprocessing process to obtain key fiber information and data enhancement to prepare for subsequent fiber feature extraction, and then perform the LBP (Local Binary Pattern) [5], texture feature and GLCM (gray-level co-occurrence matrix) [6] feature of the fiber image according to the fiber image characteristics and various feature characteristics. Extract, and then fuse multiple types of texture features into a multi-dimensional feature matrix vector. Then this paper use the LC-KSVD algorithm to establish a cashmere and wool classification model to realize the classification and identification of cashmere wool. The method in this paper realizes the transformation of the image into the feature matrix, thereby reducing the input pressure of the classification model.

2. Methods

2.1. Image preprocessing

As shown in Figure 1, when acquiring the original cashmere wool fiber image, it is necessary to process the fiber and then use a microscope for imaging, which will inevitably increase interference information, such as bubbles and visible impurities in image 1(a)(b). First, this method use median filtering and dark channel prior dehazing algorithm [7] to make the blurred original image clearer and facilitate the extraction of subsequent grayscale features, and then make the image more in line with human visual effects, and then adopt adaptive threshold binarization, morphological operation, remove excessive impurities in the image. The results are shown in Figure 2(a)(b)(c). It can be seen from Figure 2 that after the morphological operation processing, the fiber contour area in the fiber binarization map is well extracted, the adhesion between the fiber and the noise is improved, the fiber edge is well separated, and the edges of the fiber are complete, and some impurities are removed, which can better separate the fiber and the background, and provide convenience for subsequent extraction of the target object.
2.2. Obtaining multi-dimensional feature matrix

This paper will use LBP features and GLCM to achieve the extraction of cashmere wool fiber texture feature parameters, through fixed feature dimensions to ensure the unity of image information. Due to the diversity of images, the number of key points extracted by key point-based feature extraction methods such as SIFT features are different, which will cause the dimensions of the acquired feature vectors to be difficult to unify, and it is not convenient to implement the dictionary learning process. The acquisition of the multi-dimensional feature matrix is achieved through the following three steps.

1) Extract the LBP features of the image

After image preprocessing, LBP feature extraction is performed on the image. First, the image is divided into blocks, and the elements in each block are LBP encoded [8], as shown in equation (1).

This method takes the eight elements in the neighborhood of the central element \((x_c, y_c)\), if the gray value is not less than the central element, the value is 1, otherwise it is 0, and finally an eight-bit binary code is formed as the LBP value of the central element.

\[
LBP(x_c, y_c) = \sum_{i=0}^{7} s(g_i - g_c) \times 2^i
\]  

(1)

2.3. In the equation (1), \(g_c\) represents the central gray value, \(g_i\) represents the neighborhood gray value, and \(s(x)\) is the symbolic function, as shown in equation (2).

\[
s(x) = \begin{cases} 
1, & x \geq 0 \\
0, & x < 0 
\end{cases}
\]  

(2)

Then this method count the LBP value of each block to obtain the LBP feature vector of this image \(y = [a_1 \ldots a_m]\).

2) Extract the GLCM of the image

This paper will mainly use four parameters of contrast, entropy, energy and homogeneity to extract texture features of cashmere wool images. It can be seen from the literature [9] that the gray-level co-occurrence matrix is affected by the pixel pitch and the generation angle, so for the generation angle, the average value of the texture feature parameters of the four angles of 0, 45, 90, 135 is selected as the feature parameter and the pixel spacing is 4. This paper is to extract the gray-scale feature matrix from the denoised image.

3) Construct image matrix vector

After extracting the LBP feature \(a\) and the gray level co-occurrence matrix feature parameter \(b\) for each image, the two types of features are fused into feature \(y\) to form a feature matrix \(Y\) of dimension, \(Y\) represents a matrix vector with \(n\) images, as shown in equation (3) Show.

\[
Y = \{y_1 \ldots y_k \ldots y_n\}
\]  

(3)

Among them, \(y_k[k = 1,2,\ldots,n]\) is expressed as the k-th image among \(n\) images, which is expressed by a dimensional matrix (vector).

2.4. Label Consistent K-SVD algorithm

Taking into account the characteristics of the diversity of fiber images, in order to accuracy and running time and other factors, this paper uses the label consistency KSVD algorithm. The classification is realized by using the obtained classifier \(W\) and the sparse coding \(X\) of the test set, and the classification principle is shown in equation (4).

\[
c = \arg \max_c (l - WX_{test})
\]  

(4)
Among equation (4), $l$ is the actual category of the test set, and $c$ is the predicted category.

3. Experiments

3.1. Materials

The data set is mainly composed of optical microscope fiber images. The optical microscope fiber image is provided by Erdos Cashmere Group. In order to obtain more different fiber shapes, the image is pre-processed and the image is enhanced. The optical microscope data set is shown in tab. 1.

### Table 1 Experimental data set situation

| Data name | Cashmere |  | Wool |  |
|-----------|----------|---|------|---|
|           | trainset| testset | trainset | testset |
| Data size | 4607     | 1618 | 3950 | 1626 |

3.2. Experimental details

In this paper, texture feature extraction is performed according to the characteristics of fiber images, and the image matrix vector $Y$ is constructed. The original image size is 256*256. After feature extraction of a single image, the feature dimension is 1*14, and the final training set image matrix vector $Y$ is 14*8557, The image feature matrix of the test set is 14*3244. Compared with the original image directly as the data set, the method in this paper significantly reduces the matrix dimension, which is beneficial to the calculation of the subsequent classification model.

### Table 2 Classification results of different numbers of dictionaries

| Number of dictionaries | iterations | Classification time(s) | Accuracy (%) |
|------------------------|------------|------------------------|--------------|
| 500                    | 1000       | 0.090                  | 82.6         |
| 530                    | 1000       | 0.093                  | 81.3         |
| 540                    | 1000       | 0.090                  | 82.3         |
| 560                    | 1000       | 0.093                  | 84.8         |
| 570                    | 1000       | 0.094                  | 89.6         |
| 580                    | 1000       | 0.094                  | 83.5         |
| 600                    | 1000       | 0.093                  | 82.0         |

It can be seen from the literature [4] that the final classification effect of dictionary learning is mainly affected by the number of dictionaries and the number of iterations. As shown in tab. 2, this article tested different numbers of dictionaries under the same number of iterations. It can be seen that when the number of dictionaries is 570, the classification result is the best, the accuracy rate can reach 90%, and the classification speed is faster. It can also be seen that it is not that the greater the number of dictionaries, the better the classification effect. It is necessary to conduct multiple experiments on the data set to find the best parameters for the fiber image. At the same time, this paper verifies the influence of the number of iterations based on the best classification effect, that is, when the number of dictionaries is 570, the effect of different iterations is shown in tab.3. It can be seen that when the number of iterations is 1000, the accuracy is the highest, and the classification time is almost indistinguishable.

### Table 3 Classification results when the number of dictionaries is 570

| Number of dictionaries | iterations | Classification time(s) | Accuracy (%) |
|------------------------|------------|------------------------|--------------|
| 570                    | 100        | 0.093                  | 80.4         |
| 570                    | 200        | 0.093                  | 83.2         |
| 570                    | 300        | 0.090                  | 85.1         |
| 570                    | 500        | 0.092                  | 87.7         |
| 570                    | 600        | 0.094                  | 88.3         |
3.3. Comparison with other retrieval methods
In order to better express the superiority of the method in this paper, this paper compares the method with multi-class dictionary learning algorithms, and the results are shown in Table 4. The first method is to extract LBP features of fiber images and then classify them by SVM, while the second method is to use LBP features for KSVD dictionary learning\cite{10} to obtain the dictionary, and then use SVM to achieve classification, and the third is based on the SRC classification algorithm\cite{11}, the fourth is an improvement to the KSVD algorithm. The experimental results show that the texture information can be better expressed through the fusion of LBP features and GLCM features, and the method in this paper can obtain a better classification effect by using a number of dictionaries far less than 4000.

| Name of methods        | Number of dictionaries | Accuracy(%) |
|------------------------|------------------------|-------------|
| LBP+SVM                | -                      | 48.12       |
| LBP+KSVD+SVM           | -                      | 49.88       |
| SRC                    | 4000                   | 53.32       |
| LBP+KSVD+SRC           | 4500                   | 90.1        |

4. Conclusion
This paper proposes a multi-feature fusion cashmere wool classification algorithm based on the LC-KSVD algorithm, which realizes the cashmere wool classification by fusing the texture features of the multi-morph fiber image. Due to the lack of original data, data augmentation is achieved by means of data enhancement. Compared with other classifiers, the classification effect is better, but its accuracy still needs to be improved. This method can effectively reduce the number and dimensions of key features.

Acknowledgments
Thanks to the lab's classmates for their help in the paper writing process.

References
[1] Y. Zhang. “Comparison of cashmere and wool fiber fineness results detected by optical projection microscope and electronic microscope methods,” Wool Textile Journal, vol.47(03), Mar.2019, pp.77-80, doi:CNKI:SUN:MFKJ.0.2019-03-014.
[2] Y. Su and F. Hou. “Effect of decolorization on fiber identification of wool/cashmere/polyester blended fabric,” Wool Textile Journal, vol.48(6), Jun.2020, pp.43-46, doi: 10.19333/j.mfkj.20190802604.
[3] X.D. Dong, S.Y. Yang, N. Ding and Y.B. Yu. “The Research Advances of DNA Technology in Identification for Animal Fibers,” China Fiber Inspection, vol.17, Oct.2015, pp.70-72, doi:10.3969/j.issn.1671-4466.2015.17.026.
[4] Jiang Z , Zhe L , Davis L S . Learning A Discriminative Dictionary for Sparse Coding via Label Consistent K-SVD[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2011.
[5] Ojala T , Pietikainen M , Harwood D . Performance evaluation of texture measures with classification based on Kullback discrimination of distributions[C]// Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing. Proceedings of the 12th IAPR International Conference on. 1994.
[6] Jianhui Feng, Yujing Yang. Study of Texture Images Extraction Based on Gray Level Co-Ocurrence Matrix[J]. Beijing Surveying and Mapping, 2007(03):19-22. doi: 10.3969/j.issn.1007-3000.2007.03.006.

[7] K. He, J. Sun and X. Tang, “Single image haze removal using dark channel prior,” 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, Aug. 2009, pp. 1956-1963, doi: 10.1109/CVPR.2009.5206515.

[8] L.P. Yao and Z.L. Pan. “Research on face recognition method based on improved HOG and LBP algorithms,” Optoelectronic Technology, vol. 40(2), Jun. 2020, pp. 114-118, doi: 10.19453/j.cnki.1005-488x.2020.02.008.

[9] Wenyu XING, Na DENG, Binjie XIN, Chen YU. Identification of wool and cashmere based on multi-feature fusion image analysis technology[J]. Journal of Textile Research, 2019, 40(03): 146-152.

[10] M. Aharon, M. Elad and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," IEEE Transactions on Signal Processing, vol. 54(11), Nov. 2006, pp. 4311-4322, doi: 10.1109/TSP.2006.881199.

[11] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and Y. Ma, "Robust Face Recognition via Sparse Representation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31(2), Feb. 2009, pp. 210-227, doi: 10.1109/TPAMI.2008.79.