Insect Recognition Based on Complex Frequency Domain Transform and Multi-feature Fusion

Wenxia Bao*, Xiang Qiu, Linsheng Huang, Dong Liang, Yongluo Li
National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University, Hefei, Anhui230601, China

*Corresponding author’s e-mail: bwxia@ahu.edu.cn

Abstract. The accurate recognition of insects is of great significance for the protection and treatment of crops, vegetables and fruit trees. In order to improve the accuracy of insect images classification at the order level, the insect recognition algorithm based on complex frequency domain transform and multi-feature fusion is proposed in this paper. Firstly, the color features and Histogram of Oriented Gradient features of the insect images were extracted in the spatial domain. Secondly, the multi-scale and multi-directional complex frequency domain transform was performed on the images using the Multi-directional Dual Tree Complex Wavelet Transform. The shape features and texture features were obtained from the low frequency sub-band images. Finally, the spatial domain features and the complex frequency domain features were fused using the Relief-F algorithm, and the Support Vector Machine was used to recognition the insect images at the order level. The results of insect images recognition at nine orders level showed that the proposed method of multi-feature fusion in complex frequency domain and spatial domain can significantly improve the accuracy of insect recognition.

1. Introduction
Currently, there are about 900,000 species of living insects known in the world. Most scholars and experts believe that many insect species have been still undiscovered [1]. Due to the large number of insect species, it is difficult to collect data sets, and the identification of insects also requires a large number of experienced biologists or trained technicians to complete. Therefore, the classification of insects is a very labor-intensive and time-consuming job. The accurate classification of insects can predict insect diseases and protect crops and fruit trees. Therefore, the combination of modern image processing and analysis algorithm has been introduced to more accurately classify insects.

Techniques such as image processing and pattern recognition can help humans to recognize insects. Domestic and foreign scholars have done a lot of research in the field of insect recognition by computer vision technology. Hafiz Gulfam et al [1] explored the technique of classifying insects by using Bayesian networks, and classifying four different types of insects by extracting the color, shape and SIFT features of insect images. Xie Chengjun et al [2] proposed a feature extraction method combining sparse coding with spatial pyramid. In this method, the recognition rate on 35 kinds of pests was improved by about 9% comparing with the extracted color histogram, Local Binary Patterns (LBP) and shape features. Lu A. et al [3] proposed a hybrid method of Discriminating Local Soft coding (DLSoft), in which local and discriminant coding strategies were combined, and local soft coding and class-specific codebooks for discriminant representation were obtained by using adjacent codewords.
In the past, the recognition of insects was mainly at the species level, but it is also important for recognition at the order level, especially in popular science and initial insect identification. Taxonomists usually need to identify specimens to the order level before naming species. However, because of the high order, insects in different orders will have similar color shapes, which makes it difficult for public and primary taxonomists to distinguish. Therefore, the recognition of insects at the order level is an important part of the entire insect recognition process [4]. In most insect recognition algorithms, features were extracted in the spatial domain, which might contain a lot of redundant information due to the complex spatial frequency of the actual image. Therefore, the extracted features might not be very effective. In this paper, for the insect recognition at the order level, M-DTCWT was used to transform the image into complex frequency domain, and effective shape features and texture features were extracted in the low frequency sub-band image. Besides, the spatial domain features are obtained, and multi-feature fusion was realized by Relief-F algorithm, which improved the recognition accuracy of insect image.

2. Materials and methods

2.1. Data collection
In this paper, 225 sample images of 9 insect orders were selected with 25 of each order [12]. The resolution of the images was unified to 150 pixels × 150 pixels. The orders and numbers of 9 insect orders are listed in Table 1. Some sample insect images are showed in Figure 2. 153 images were randomly selected as training samples, and 72 were used as test samples. The experiments were performed on Matlab2016a.

| Label | Order          | Number of pictures | Label | Order          | Number of pictures |
|-------|----------------|--------------------|-------|----------------|--------------------|
| 1     | Auchenorrhyncha| 25                 | 6     | Megaloptera    | 25                 |
| 2     | Coleoptera     | 25                 | 7     | Neuroptera     | 25                 |
| 3     | Heteroptera    | 25                 | 8     | Odonata        | 25                 |
| 4     | Hymenoptera    | 25                 | 9     | Orthoptera     | 25                 |
| 5     | Lepidoptera    | 25                 |       |                |                    |

Figure 1 Insect sample images.

2.2. Feature extraction and selection

2.2.1. Spatial domain feature
The color feature can be used to describe the surface properties of the scene corresponding to the image or image area [7]. The HSI color model can better reflect people's perception and discrimination of color. At the same time, H and S components can reduce the influence of illumination changes.

The Histogram of Oriented Gradient can effectively represent image information in complex environments, which extracts information about the distribution of the target edges in the local area as a means of representing the shape of the target [8].

Horizontal and vertical gradient values are as follows:

\[ G_x(x, y) = H(x + 1, y) - H(x - 1, y) \]  \hspace{1cm} (1)

\[ G_y(x, y) = H(x, y + 1) - H(x, y - 1) \]  \hspace{1cm} (2)
where $H(x, y)$ represents the pixel value of a point. $G_x(x, y)$ and $G_y(x, y)$ respectively determine the horizontal direction gradient and the vertical direction gradient at the pixel point $(x, y)$ in the input image.

The gradient magnitude and direction of the pixel are as follows:

$$G(x, y) = G_x(x, y) - G_y(x, y)$$

$$\alpha (x, y) = \tan^{-1} \frac{G_x(x, y)}{G_y(x, y)}$$

where $G(x, y)$ and $\alpha$ calculate the gradient amplitude and direction at the pixel $(x, y)$ respectively.

2.2.2. Complex frequency domain feature

Image information on the spatial domain is more complicated, and feature extracted directly contains a lot of redundant information. Therefore, it is necessary to find a method that transforms the original image into different components, and the features of the image can be extracted from these components. The redundant information can be avoided [13].

The Dual Tree Complex Wavelet Transform (DTCWT) [11] has invariant translation, better direction selectivity and finite redundancy. Besides, it can increase the directional selectivity of DTCWT, combined with the filter bank and the hourglass filter bank in DTCWT which can conclude Multi-direction Dual-Tree Complex Wavelet Transform (M-DTCWT) filter bank. Comparing with DTCWT, M-DTCWT can more effectively represent the geometric properties of shapes and textures with multi-directionality in two-dimensional images.

![Diagram of M-DTCWT decomposition in an image.](diagram.png)

Figure 2. M-DTCWT decomposition in an image.

Let the complex scale function be $\Phi_{i,j}(n,n)=\Phi^c_{i,j}(n,n)+\sqrt{-1}\Phi^i_{i,j}(n,n)$, then we can represent the complex wavelet function as $\Psi_{i,j}(n,n)=\Psi^c_{i,j}(n,n)+\sqrt{-1}\Psi^i_{i,j}(n,n)$, where $n_1$, $n_2$ are the corresponding points on the image, $j, k$ are the expansion and translation index, $i$ is the number of sub-bands in the direction and $\text{re, im}$ are the real and imaginary parts after decomposition. M-DTCWT decomposes the two-dimensional image $f(n_1,n_2)$ showed in Figure 3 through a series of complex scaling functions and complex wavelet functions:

$$f(n_1,n_2) = \sum_{k \in Z} c_{j,k} \Phi_{i,k} (n_1,n_2) + \sum_{j, k} \sum_{i, a, b} d_{j,k}^{i,a} \Psi_{i,j}^{a,b} (n_1,n_2)$$

where $Z$ represents the natural number set, $c_{j,k} = \left\langle MS_{i,j}^{a,b}(n_1,n_2), \Phi_{i,j}^{a,b}(n_1,n_2) \right\rangle$ represents the scale factor, and $d_{j,k}^{i,a} = \left\langle MS_{i,j}^{a,b}(n_1,n_2), \Psi_{i,j}^{a,b}(n_1,n_2) \right\rangle$ represents the complex wavelet coefficients in the $i$-th direction.

Most of the information in the image is concentrated in the low frequency, and the high frequency information still retains a lot of redundant information. The details feature of shape or texture information are characterized by gradual changes. In the classical wavelet transform method, each layer of decomposition space can only provide direction information of three angles: horizontal, vertical and oblique. And it is often difficult to adapt to continuous changes in direction. The dual tree complex wavelet transform provides eight alternative directions: two low frequency sub-bands and six high frequency sub-bands, and only very limited redundancy information. For two-dimensional image
signals, the redundant information is independent of the number of scaled spatial layers and remains at 1:4. In order to ensure that the low-frequency information is not lost, the average of the two low-frequency sub-bands is taken as low-frequency information, in which the shape and texture features are extracted. Hu invariant moments are used to extract shape features of low frequency images.

The shape information of objects in the image can be better described by Hu invariant moments. The second- and third-order normalized central moments are used to construct seven invariant characteristics of the image. The constant rotation, translation and scaling features of continuous image processing are described by Hu invariant moment.

The texture reflects the visual characteristics of homomorphism in the image, which is represented by the grayscale distribution of the neighborhood of the pixel and its surrounding space. Texture can reduce the difference of eigenvalues of samples in the class, and increase the difference of eigenvalues between sample classes. The Gray Level Co-occurrence Matrix (GLCM) was extracted as the texture information of insect images in this paper, which can be used to obtain statistical information about the distribution of pixel pairs. “Direction” and “distance” was defined to calculate GLCM. It is possible to analyze the pixel pairs by them and then calculate the number of pixel pairs with a certain gray value distribution [6]. GLCM is used to count the pairs of the pixel with gray value \( i \) and \( j \) at the same time. A distance \( \sigma \) of 1 is calculated for each insect image, and a co-occurrence matrix of 4 directions is generated.

Through multi-directional dual tree complex wavelet decomposition of gray image, low-frequency image was extracted in complex frequency domain and divided into blocks. Then the mean values of four texture feature statistics, correlation, energy, homogeneity and contrast, were taken as insect texture features. The shape features of the insect images were represented by 7 Hu invariant moments.

2.3. Multi-feature fusion based on Relief-F

The Relief-F algorithm is an important attribute weighting method based on classification labels [9].

The correlation of each feature and category can be calculated according to Relief-F algorithm. When the weight of a feature is less than a certain threshold, the feature can be removed. A sample \( T \) is randomly selected from the training sample set, then \( k \) neighbor samples of sample \( T \) are found from each class different \( T \). Then the weight of each feature is updated.

The calculation of feature weights is as follows:

\[
W(A) = W(A) - \sum_{j=1}^{2} \text{diff} \left( A, R_{j}, H_{j} \right) / (mk) \\
+ \sum_{\text{class}(R)} \left[ \frac{p(C)}{1-p(\text{class}(R))} \right] \sum_{j=1}^{2} \text{diff} \left( A, R_{j}, M_{j}(C) \right) / (mk)
\]

(6)

where \( \text{diff}(A,R,R_{j}) \) represents the difference between the samples \( R \) and \( R_{j} \) in the feature vector \( A \). \( m \) represents the number of samples, \( \text{class}(R) \) is the \( R \) class, \( p(C) \) represents the probability of class \( C \) of \( R \), \( p(\text{class}(R)) \) represents the probability of class \( R \), and \( M_{j}(C) \) represents the \( j \)-th nearest neighbor sample in class \( C \). \( \text{diff}(A,R,R_{j}) \) is expressed as follows:

\[
\text{diff} \left( A, R_{1}, R_{2} \right) = \begin{cases} 
\left| \frac{A_{1} - R_{1}}{\max(A) - \min(A)} \right| & \text{When } A \text{ is continuous} \\
0 & \text{When } A \text{ is discrete or } R_{1} \neq R_{2} \\
1 & \text{When } A \text{ is discrete or } R_{1} = R_{2}
\end{cases}
\]

(7)

In this paper, the color features and HOG features of the insect image were extracted in the spatial domain, and the shape features and texture features were extracted in the complex frequency domain for concatenation fusion to obtain the post-fusion feature set. Different weights to the features according to the relevance of the corresponding categories of the fused feature sets were assigned by Relief-F algorithm. By setting the threshold to 0.01, the weight of each feature was calculated. When
the calculated weight of the feature was less than 0.01, the feature is removed, and when the feature weight was greater than or equal to 0.01, the obtained feature was reconcatenated to form a new feature set. The new feature set was put into the classifier for classification.

2.4. Algorithm flow
The flow of the insect recognition algorithm based on multi-feature fusion of complex frequency domain is shown in Figure 4. The HOG and color features were extracted on the images of the processed training set. Then M-DTCWT method was used to decompose images in multi-scale and multi-direction to obtain high and low frequency coefficients. And then the average of the two low-frequency coefficients was taken as low-frequency information, in which the shape features and texture features were extracted respectively. Finally, the weights of different features were calculated according to the correlation of each feature by using Relief-F, and the images were classified by Support Vector Machine (SVM) classifier [10].

![Figure 3. Algorithm flow chart](image)

3. Results and discussion

3.1. Experimental result
M-DTCWT method was used to decompose the insect images in multi-scale and multi-direction. The accuracy of insect image recognition in different decomposition layers was given in Figure 4. After only one decomposition, the best recognition result can be obtained. Because the first layer can not only retain most of the information of the original insect image, but also filter a lot of redundant information. After several layers decomposition, many information in insect images was filtered out, resulting in a gradual reduction in recognition rate.

![Figure 4. Recognition accuracy results of different decomposition levels.](image)
In order to verify the validity of feature extraction for individual insect order classification, SVM was used to classify under different feature sets (Fig. 5). The spatial domain features 1 represent the color features and HOG features, which were extracted in the spatial domain. The spatial domain features 2 represent the color features, shape features, texture features and HOG features, which were extracted in the spatial domain, too. In the method of this paper, the color features and HOG features were extracted in the spatial domain while the texture features and shape features were extracted under the complex frequency domain.

Then the different classification methods were compared. The recognition results of SVM, Logistic Regression (LR), K-Nearest Neighbor (KNN), and Random Forest (RF) in different feature sets are shown in Table 2.

| Feature set                | Algorithm | LR     | KNN   | RF    | SVM   |
|----------------------------|-----------|--------|-------|-------|-------|
| spatial domain feature 1   | 79.1      | 61.4   | 70.3  | 80.6  |
| spatial domain feature 2   | 79.1      | 77.8   | 81.9  | 87.5  |
| Method of this paper       | 90.2      | 85.1   | 87.0  | 97.22 |

3.2. Analysis of results

Compared with different features, our method achieved the best results in the classification of individual insect order, especially insect orders of Auchenorrhyncha, Heteroptera, Neuroptera and Odonata (Fig. 5). Compared with the methods of LR, KNN, and RF, the SVM classifier was superior in multi-feature multi-classification under the same feature set (Table 2). Compared with the methods of the three papers are listed in Table 3, the method was proposed in this paper had a better recognition rate of 97.2%.

4. Conclusion

The insect recognition algorithm was proposed based on complex frequency domain transform and multi-feature fusion in this paper. The M-DTCWT transform method was introduced to decompose the image into high frequency coefficients and low frequency coefficients. The low-frequency shape features and texture features of the images were extracted in the complex frequency domain. At the same time, the color features and HOG features of the insect images were obtained in the spatial domain. Then the Relief-F algorithm and SVM method were used for insect recognition. Through the comparison of experiments, the effectiveness of the proposed method for different insect orders recognition is verified.
Acknowledgments
This study was funded by National Key Research and Development Program of China (2016YFD0800904), National Natural Science Foundation of China (41771463, 61672032), and Anhui Provincial Science and Technology Project (16030701091).

References
[1] Umar, H.G.A., Abbas, Q., Gulzar, F. (2017) Insect classification using image processing and bayesian network. Journal of Entomology and Zoology Studies, 5(6):1079-1082.
[2] Xie, C.J., Li, R., Dong, W., et al. (2016) Recognition for insects via spatial pyramid model using sparse coding. Transactions of the CSAFE, 32(17):144–151.
[3] Lu, A., Hou, X., Liu, C. L., and Chen, X. (2012). Insect species recognition using discriminative local soft coding. International Conference on Pattern Recognition. IEEE. Tsukuba, 1221–1224.
[4] Chen, X.L., Hou, X.W., Liu, C.L., et al. (2008) Advances in the automated insect image identification. Chinese Bulletin of Entomology, 45(2), 317-322.
[5] Shao, J., Pang, X.Y., Yang Z.J., Li, J.L. (2017) Research on intelligent recognition of axis orbit based on Hu moment invariants and fractal box dimension. 2017 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). Jeju, pp. 794-799.
[6] Shang, Z. and Li, M. (2016) Combined Feature Extraction and Selection in Texture Analysis. (2016) 2016 9th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou. pp. 398-401.
[7] Pydipati, R., Burks, T.F., and Lee, W.S. (2006) Identification of citrus disease using color texture features and discriminant analysis. Computers and Electronics in Agriculture. 52(1-2), 49-59.
[8] Korkmaz, S.A., Akçıçek, A., Bînol, H. and Korkmaz, M.F. (2017) Recognition of the stomach cancer images with probabilistic HOG feature vector histograms by using HOG features. 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY), Subotica, pp. 000339-000342.
[9] Zhang, L., Li, D. and Zhong, C. (2009) Collaborative Optimization of Clustering by Fuzzy c-means and Weight Determination by ReliefF. 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery, Tianjin. pp. 454-459.
[10] Chrystal, J.B., and Joseph, S. (2015) Multi-label classification of product reviews using structured SVM. International Journal of Artificial Intelligence & Applications. 3: 61-68.
[11] Zhang, X. and Zhang, R. (2012) The technology research in decomposition and reconstruction of image based on two-dimensional wavelet transform. 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery, Sichuan. pp. 1998-2000.
[12] Wang J., Lin C. Ji L., Liang A. (2012) A new automatic identification system of insect images at the order level. Knowledge-Based Systems. 33, 102---110.
[13] Yang, P., Zhang, F., Yang, G. (2018) Fusing dctw and lbp based features for rotation, illumination and scale invariant texture classification. IEEE Access, PP (99), 1-1.