Insect Species Recognition using Sparse Representation

An Lu Student¹ alu@nlpr.ia.ac.cn
Xinwen Hou Prof¹ xwhou@nlpr.ia.ac.cn
Xiaolin Chen Prof² xlchen@ioz.ac.cn
Chenglin Liu Prof¹ liucl@nlpr.ia.ac.cn

¹ Institute of Automation Chinese Academy of Sciences Beijing, China

² Institute of Zoology Chinese Academy of Sciences Beijing, China

Insect species recognition is a typical application of image categorization and object recognition. In this paper, we propose an insect species recognition method based on class specific sparse representation. On obtaining the vector representation of image via sparse coding of patches, an SVM classifier is used to classify the image into species. We propose two class specific sparse representation methods under weakly supervised learning to discriminate insect species which have substantial similarity to each other. Experimental results show that the proposed methods perform well in insect species recognition and outperform the state-of-the-art methods on generic image categorization.

Our work is motivated by the sparse coding spatial pyramid matching (ScSPM) model for generic image categorization [3], sparse representation based classification (SRC) algorithm for face recognition [2] and image restoration method [1]. All of the three methods are based on sparse representation. Comparing to these methods, we calculate bases of each class to obtain the class specific sparse representations in the strategies of minimal reconstruction residual and sparsity of local features.

Our bases construction method is based on [3]. However their method is more suitable for generic object image datasets which have more distinction between classes than that of insect species. So motivated by the work of [2], we adopt a class specific bases construction strategy. The holistic optimization problem can be formulated as:

$$\min_{B_{i},S^{(i)}} \sum_{n_{i}=1}^{N_{i}} \|x_{n_{i}}^{(i)} - B_{i}s_{n_{i}}^{(i)}\|_{2}^{2} + \lambda \|s_{n_{i}}^{(i)}\|_{1}, \text{ for } i = 1,2,\dots,C$$

$$s.t. \|b_{i,k_{i}}\|_{2} \le c, \forall k_{i} = 1,2,\dots,K_{i}$$
(1)

For each class we calculate its own basis matrix by an iterative process. Firstly we randomly initialize basis matrix B_i to calculate the new sparse codes $s_{n_i}^{(i)}$ for input vector $x_{n_i}^{(i)}$ for each class i:

$$\min_{s_{n_i}^{(i)}} \|x_{n_i}^{(i)} - B_i s_{n_i}^{(i)}\|_2^2 + \lambda \|s_{n_i}^{(i)}\|_1$$
 (2)

Then we fix sparse codes S and solve the following optimization problem with constraint:

$$\min_{B_i} \sum_{n_i=1}^{N_i} \|x_{n_i}^{(i)} - B_i s_{n_i}^{(i)}\|_2^2
s.t. \|b_{i,k_i}\|_2 \le c, \ \forall k_i = 1, 2, \dots, K_i$$
(3)

At the same time, for any new input vector x_{new} , we can get C coefficient vectors (sparse codes) $s_{new}^{(i)}$ respectively to each basis matrix by solving the optimization problem:

$$\min_{(i)} \|x_{new} - B_i s_{new}^{(i)}\|_2^2 + \lambda \|s_{new}^{(i)}\|_1, \quad for \quad i = 1, 2, \dots, C$$
 (4)

We proposed two strategies to concatenate the C coefficient vectors into one sparse vector to represent the original input vector. The first one we call it minimal residual class specific sparse representation (MRCSSR). That means we take the coefficient vector which minimizes the residual of reconstruction as its original value and other vector as zero.

$$p = \arg\min_{i} \|x_{new} - Bs_{new}^{(i)}\|_{2}$$

$$s_{new} = [\underbrace{0, 0, \dots, 0}_{K_{1}}, \underbrace{0, 0, \dots, 0}_{K_{2}}, \dots, \underbrace{(s_{new}^{(p)})^{T}}_{K_{D}}, \dots, \underbrace{0, 0, \dots, 0}_{K_{C}}]^{T}$$
(5)

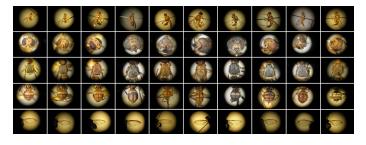


Figure 1: Example of our Tephritidae dataset

Sub-dataset	Whole	Head	Thorax	Abdomen	Wing
Species	19	19	20	17	14
Images	152	143	151	144	103

Table 1: Number of species and images in each sub-dataset.

The second strategy we call it sparsest class specific sparse representation (SCSSR). That means we take the coefficient vector which is sparsest to represent the original feature and other vector as zero. Here we take L_0 norm to evaluate the sparsity of the coefficient vectors.

$$p = \arg\min_{i} \|s_{new}^{(i)}\|_{0}$$

$$s_{new} = \underbrace{[0, 0, \dots, 0, \underbrace{0, 0, \dots, 0}_{K_{2}}, \dots, \underbrace{(s_{new}^{(p)})}^{T}, \dots, \underbrace{0, 0, \dots, 0}_{K_{C}}]^{T}}_{K_{C}}$$
(6)

After calculating the sparse representation of each input vectors, any pooling method such as averaging or max pooling [3] can combine these sparse codes of input vectors belonging to the same sample together to obtain the final feature vectors. Then any learning method such as neural networks or SVM is competent for the recognition task.

Our Tephritidae dataset is composed of 3 genera and 20 species. Each specimen is taken one photograph respectively of its whole body, head, thorax, abdomen and wing (as shown in Fig.1). So we divide the whole dataset into 5 sub-dataset according to different part of specimen. Table. 1 shows the number of species and photographs of the 5 sub-datasets.

Our experiments on the Tephritidae dataset and Caltech101 dataset demonstrated the effectiveness of our methods. We believe that constructing a basis matrix for each class will take more discriminative information into the final sparse representation and both the minimal residual and the sparsest strategy remove some noise among other similar classes and remain the information which is utmost expressive for the true classes.

- [1] J. Mairal, M. Elad, and G. Sapiro. Sparse representation for color image restoration. *IEEE Trans. IP*, 17(1):53–69, 2008.
- [2] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Trans. PAMI*, 31(2): 210–227, 2009.
- [3] J. Yang, K. Yu, Y.H. Gong, and T. Huang. Linear spatial pyramid matching using sparse coding for image classification. In CVPR, 2009.