

## 基于稀疏编码金字塔模型的农田害虫图像识别

谢成军<sup>1</sup>, 李 瑞<sup>1\*</sup>, 董 伟<sup>2</sup>, 宋良图<sup>1</sup>, 张 洁<sup>1</sup>, 陈红波<sup>1</sup>, 陈天娇<sup>1</sup>

(1. 中国科学院合肥智能机械研究所, 合肥 230031; 2. 安徽省农业科学院农业经济与信息研究所, 合肥 230031)

**摘 要:** 相较于一般物体的图像, 农作物害虫图像因具有复杂的农田环境背景, 分类与识别更加困难。为提高害虫图像识别的准确率, 该文提出一种基于图像稀疏编码与空间金字塔模型相结合的害虫图像表示与识别方法。该方法利用大量非标注的自然图像块构造完备学习字典, 并运用该学习字典实现对害虫图像的多空间稀疏表示。与此同时, 结合多核学习, 该文设计了一种害虫图像识别算法。通过对 35 种害虫的识别, 试验结果表明: 在相同方法下, 该文所提特征提取方法可使平均识别精度提高 9.5 百分点; 此外, 进一步通过对 221 种昆虫及 20 种蝴蝶的识别, 试验结果表明: 与传统方法相比较, 该文所提方法使得平均识别精度提高 14.1 百分点。

**关键词:** 图像识别; 算法; 害虫控制; 字典学习; 稀疏编码; 金字塔模型

doi: 10.11975/j.issn.1002-6819.2016.17.020

中图分类号: TP751

文献标志码: A

文章编号: 1002-6819(2016)-17-0144-08

谢成军, 李 瑞, 董 伟, 宋良图, 张 洁, 陈红波, 陈天娇. 基于稀疏编码金字塔模型的农田害虫图像识别[J]. 农业工程学报, 2016, 32(17): 144—151. doi: 10.11975/j.issn.1002-6819.2016.17.020 http://www.tcsae.org

Xie Chengjun, Li Rui, Dong Wei, Song Liangtu, Zhang Jie, Chen Hongbo, Chen Tianjiao. Recognition for insects via spatial pyramid model using sparse coding[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(17): 144—151. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2016.17.020 http://www.tcsae.org

### 0 引 言

目前全世界昆虫约 1 000 万种, 已经发现并且正式命名的约 100 万种, 其中包括农作物害虫。中国是一个农业大国, 害虫种类繁多。2010 年至 2014 年间, 全国农业技术推广服务中心联合国内 42 家植保领域相关单位调查表明, 仅水稻、小麦、玉米等 15 种主要农作物的害虫就达 2000 余种<sup>[1]</sup>。对于农作物害虫的识别, 传统的方法是植保专家依据害虫的特征进行人工识别, 诊断的准确率和专家的知识层次、经验水平以及主观意识密切相关, 存在一定的主观性和局限性。此外, 现行的害虫分类、识别工作主要是依靠少数有经验的植保专家来完成, 但害虫种类繁多, 每一位植保专家穷其所能, 也只能识别部分害虫。近些年, 计算机视觉技术和理论飞速发展, 在农业、工业等方面发挥了重要作用, 同时也为害虫的自动分类与识别提供了可靠的技术手段。利用其快捷、方便、准确、客观的优点, 改造传统害虫分类与识别模式, 使其最终能够部分代替领域植保专家对农作物害虫进行有效及时地自动化识别, 具有非常重要的研究价值

和现实意义。

近几十年, 基于计算机视觉技术的害虫图像自动分类与识别得到了广泛的研究。中国农业大学沈佐锐教授较早将计算机视觉技术运用到昆虫图像识别<sup>[2-4]</sup>中, 他率先将数学形态学作为图像特征引入到机器学习中, 并运用到昆虫分类上, 获得了较好的分类效果。杨红珍等<sup>[5]</sup>通过提取害虫图像中的颜色、形状等特征, 并使用径向基神经网络算法对提取的特征进行训练学习, 得到了害虫分类模型, 从而建立了害虫远程自动识别系统。基于支持向量机(support vector machine, SVM)-模拟退火算法, 胡玉霞等<sup>[6]</sup>提出了储粮害虫自动分类方法。在害虫预测预报方面, 毛文华等<sup>[7]</sup>以新疆、青海、甘肃等地区的蝗虫为研究对象, 提出了基于机器视觉技术的蝗虫自动识别方法, 获得了较好的效果。在储粮害虫自动识别方面, 张红涛等<sup>[8]</sup>通过提取储粮害虫图像 17 个形态学特征, 采用支持向量机训练学习后, 在实验室环境下, 针对 9 类常见的储粮害虫进行识别试验, 识别正确率达 95% 以上。竺乐庆等<sup>[9-10]</sup>以鳞翅昆虫为例, 采用双树复小波提取图像纹理, 提出基于颜色特征和纹理特征匹配的昆虫识别自动方法, 在自然背景下, 该算法识别正确率达 76%。Wang 等<sup>[11]</sup>通过将人工神经网络与支持向量机结合, 实现害虫图像自动识别, 识别正确率达到 93% 以上。在果蔬害虫图像识别方面, Wen 等<sup>[12]</sup>尝试采用全局特征模型与局部特征模型相结合的特征提取方式获得害虫图像特征, 试验结果表明在运用联合特征提取模式情况下, 果树害虫图像识别正确率达 86.6%。为实现田间小麦蚜虫自动计数与识别, Liu 等<sup>[13]</sup>通过提取小麦蚜虫图像 HOG 特征并支持向量机训练学习, 实现简单麦田背景下小麦蚜

收稿日期: 2016-02-26 修订日期: 2016-05-31

基金项目: 国家自然科学基金项目: 基于上下文感知与稀疏表示的害虫图像识别研究(31401293); 安徽省农业科学院院长青年创新基金项目: 基于机器视觉的植保图像采集与元数据管理技术研究(14B1461)。

作者简介: 谢成军, 男, 安徽全椒人, 副研究员, 博士, 主要从事机器视觉、精准农业、农业物联网等研究。合肥 中科院合肥智能机械研究所, 230031。Email: cjxie@iim.ac.cn

\*通信作者: 李 瑞, 男, 河北衡水人, 硕士, 主要从事图像处理、模式识别等方面研究。合肥 中科院合肥智能机械研究所, 230031。

Email: li\_rui1985@163.com

虫的自动计数与识别，平均识别精度达到 86.81%。

如何有效地表示图像是图像处理、识别与理解的基础，而高效的图像表示又要求用较少的信息（如图像压缩、特征抽取等）来描述图像的重要特征。近年来，越来越多的研究都集中在图像的稀疏表示方面，图像的稀疏编码也成为研究热点。稀疏编码技术<sup>[14]</sup>的核心思想是：首先通过大量的图像样本块构建过完备学习字典；其次，对于新的图像，可以通过很少的几个非零元素或很少的几个远大于零的元素，由过完备字典来线性表示或重构，其表示的系数是稀疏的。韩安太等<sup>[15-16]</sup>将稀疏表示思想引入害虫识别，利用害虫图像构造字典矩阵，通过求解 L1 范数下最优化问题，实现害虫测试样本的重构；通过对害虫图像残差结果的分析，实现害虫分类识别。胡永强等<sup>[17]</sup>将稀疏编码提取害虫图像特征，并应用 AdaBoost 算法学习不同特征下害虫的分类结果，最终实现了青海地区油菜常见害虫自动识别。此外，谢成军等<sup>[18]</sup>通过对害虫图像的多任务稀疏表示来完成图像特征提取，并结合多核机器学习方法完成分类模型训练，试验结果也验证了该方法的有效性。与谢成军等所提方法不同点在于，本文方法采用多空间稀疏编码方式，实现了不同空间下害虫图像稀疏编码。其次，引入了空间金字塔结构模型，即考虑到了害虫图像整体特征，又兼顾害虫图像局部特征，害虫图像外观特征表示更健壮。

上述害虫识别方法在实验室可控环境下建立的数据集上性能卓越。然而，在农田不可控环境下，由于图像背景、光线环境、害虫姿态等往往具有强烈的变化，因此，建立的数据集与前者有较大的差别，上述识别方法的精度以及性能就会有一定程度的下降。与传统 SIFT、HOG 等手动设计的特征提取方式不同，考虑到稀疏编码方法在图像表示方面的鲁棒性<sup>[19]</sup>，该文首先通过非监督特征学习<sup>[20]</sup>方式来获取害虫图像底层特征学习字典；其次，结合图像空间金字塔模型，提出害虫图像位置空间稀疏编码与方向空间稀疏编码，进一步拓展了害虫图像特征表达的深度与广度；最终，结合多核学习方法，提出了一种基于多核学习的害虫图像识别方法。

## 1 稀疏编码理论

近年来，越来越多的研究都集中在信号的稀疏表示方面，图像作为一种二维的信号，图像的编码表示也成为一个研究热点。在信号处理中，信号  $y \in \mathbb{R}^n$  常常可以由一组基本信号的线性组合来表示。即有如下公式

$$y = D_{n \times k} \alpha, k > n. \quad (1)$$

$D = \{d_1, d_2, \dots, d_k\}$  被称作过完备字典，字典中的每列  $\{d_i\}_{i=1}^k$  被称作原子。向量  $\alpha \in \mathbb{R}^k$  为信号  $y$  的表示系数。假设矩阵  $D$  满秩，上述公式方程组将有无穷多个解，这些解构成一个解空间。但是在实际应用中，这样的解空间并不是我们所期望得到的，信号的稀疏表示，它意欲用尽可能少的非 0 系数表示信号的主要信息（如图 1 所示），从而简化信号处理问题的求解过程。为了能够选择出带有这种特性的解，问题转化为求解以下方程

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_0 \quad \text{s.t.} \quad \|D\alpha - y\|_2 \leq \varepsilon. \quad (2)$$

式中  $\|\cdot\|_0$  是计算非零元素数量的  $l_0$  范式， $\|\cdot\|_2$  是  $l_2$  范式，参数  $\varepsilon$  定义了重构精度。由于解式 (2) 是一个 NP 问题，因此用  $l_1$  范式来取代  $l_0$  来解上面问题，并做如下定义

$$\hat{\alpha} = \arg \min_{\alpha} \|\alpha\|_1 \quad \text{s.t.} \quad \|D\alpha - y\|_2 \leq \varepsilon. \quad (3)$$

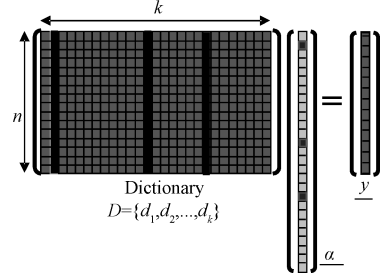


图 1 稀疏编码

Fig.1 Sparse coding

## 2 基于稀疏编码金字塔的害虫图像识别

### 2.1 害虫图像字典学习

为了学习获得一个过完备字典，首先从大量样本图像中随机抽取  $M$  个样本图像块  $y_i$ ,  $y_i \in \mathbb{R}^n$  ( $i = 1, 2, \dots, M$ )。以类似文献[20]方法，对收集来的每个图像块进行亮度、对比度以及白化等步骤的预处理。然后应用非监督学习算法创建过完备字典  $D = [d_1, d_2, \dots, d_k] \in \mathbb{R}^{n \times M}$ ,  $M$  是字典大小。为获得该学习字典，应用奇异值分解算法 (K-SVD)<sup>[21]</sup>，解最优化方程

$$\min_{D, A} \sum_{i=1}^M \|D\alpha_i - y_i\|_2^2, \quad \text{s.t.} \quad \|\alpha_i\|_0 \leq k, \forall i. \quad (4)$$

$\alpha_i \in \mathbb{R}^M$  是图像块  $y_i$  的稀疏编码， $\|\alpha_i\|_0$  表示  $\alpha_i$  中非零元素个数。

因此通过解式 (4) 最优化问题，可以得出害虫图像的过完备训练字典  $D$ 。

### 2.2 害虫图像的多空间稀疏编码

#### 2.2.1 害虫图像稀疏编码

在特征提取方面，经典的方法包括 SIFT、HOG、直方图等，一般而言，这些特征提取方法需要根据实际情况，通过人工设计并调解相应参数才能获得更好的图像识别效果。然而，在自然环境下，对于外观表征变化较大的农田害虫图像而言，依靠传统的特征提取方法解决害虫图像识别，困难显而易见。鉴于此，受到非监督特征学习的启发，本文考虑通过多空间稀疏编码来对害虫图像进行特征提取，得到更为健壮的害虫图像表示方法，从而提高识别准确率。

假设给定一幅害虫图像  $Y$ ，被分为  $N$  个局部图像块， $y_i \in \mathbb{R}^n$  ( $i = 1, 2, \dots, N$ ) 表示第  $i$  个害虫局部图像块， $D = [d_1, d_2, \dots, d_k] \in \mathbb{R}^{n \times M}$  表示给定的图像过完备学习字典。应用正交匹配追踪算法<sup>[22]</sup> (OMP-k)，每个  $y_i$  对应的稀疏编码  $\hat{\alpha}_i \in \mathbb{R}^{M \times 1}$  按式 (5) 计算

$$\hat{a}_i = \arg \min_{a_i} \|a_i\| \quad \text{s.t.} \quad \|D a_i - y_i\|_2 \leq \varepsilon. \quad (5)$$

因此, 一幅害虫图像  $Y$  的稀疏编码可表示为  $A = [a_1, a_2, \dots, a_N] \in \mathbb{R}^{M \times N}$ 。

### 2.2.2 多空间害虫图像稀疏编码

为了进一步研究害虫图像在不同空间下稀疏编码特征的分布情况, 提高害虫图像特征表示的健壮性, 本文从不同角度定义 2 种空间模式的稀疏编码: 位置空间和方向空间。

对于位置空间稀疏编码分布特征定义如下

$$f_j^l = C_l \cdot \sum_{l_i \in S_l} K\left(\left\|\frac{l_i - l_c}{h_l}\right\|\right) |a_{i,j}|. \quad (6)$$

式中  $l_i$  表示图像块的欧式坐标点,  $l_c$  表示图像块中心点坐标点,  $a_{i,j}$  表示第  $i$  个图像块的第  $j$  个稀疏系数,  $|a_{i,j}|$  表示稀疏系数不为 0 的个数,  $h_l$  是样本块窗口移动步长,  $S_l$  定义样本块窗口的欧式空间,  $C_l$  是为了规范化而设的一个常量,  $K(\cdot)$  是一个各向异性核函数。至此, 对于一个害虫图像块  $y_i$ , 位置空间稀疏编码为  $\hat{f}_i^l = [f_1^l, f_2^l, \dots, f_M^l]$ 。

因此, 一幅害虫图像  $Y$  的位置空间稀疏编码可表示为  $F^l = [\hat{f}_1^l, \hat{f}_2^l, \dots, \hat{f}_N^l] \in \mathbb{R}^{M \times N}$ 。

对于方向空间稀疏编码分布特征定义如下

$$f_j^o = C_o \cdot \sum_{o_i \in S_o} K\left(\left\|\frac{o_i - o_c}{h_o}\right\|\right) |a_{i,j}|, \quad (7)$$

$$S_l \rightarrow S_o: o = \text{atan}\left(\frac{y - y_c}{x - x_c}\right). \quad (8)$$

式中  $[x, y]$  表示图像块的欧式坐标点,  $[x_c, y_c]$  表示图像块中心点坐标点,  $S_o$  样本块窗口的方向空间,  $h_o$  是样本块窗口方向移动步长,  $C_o$  是为了规范化而设的一个常量。至此, 对于一个害虫图像块  $y_i$ , 方向空间稀疏编码为  $\hat{f}_i^o = [f_1^o, f_2^o, \dots, f_M^o]$ 。因此, 一幅害虫图像  $Y$  的位置方向稀疏编码可表示为  $F^o = [\hat{f}_1^o, \hat{f}_2^o, \dots, \hat{f}_N^o] \in \mathbb{R}^{M \times N}$ 。

### 2.3 害虫图像稀疏编码金字塔模型

在自然环境下害虫图像展现出不同表现、姿态、尺度以及发生部分遮挡等情况。常见的图像特征表示, 如颜色直方图、SIFT 等从图像整体出发很难解决图像局部遮挡、变形等问题; 从图像局部特征出发考虑的局部纹理等特征, 在图像识别时又会受图像整体表现变化影响较大。针对这些问题, 在图像表示方面即考虑到图像整体特征又兼顾图像局部信息。为此, 本文引入空间金字塔模型来解决上述问题。如图 2 所示, 本文使用 3 层金字塔来为害虫图像特征建模。在图像金字塔顶层将一幅害虫图像分为 4 个小图像块, 每个图像块经过方向空间稀疏编码以和位置空间稀疏编码, 计算后组成该层特征向量。在金字塔中间层将害虫图像依次分为互不交错的 16 个小图像块, 以同样方式计算多空间稀疏编码作为该层特征向量。在金字塔底层, 进一步将图像分为 64 个子图像块, 并以同样方式计算该层特征向量。每层图像块

的数量可以通过  $p=4^{r-1}$  ( $r=2,3,4$ ) 来计算。针对该文图像大小为  $200 \times 200$ , 金字塔顶层 4 个图像块, 每个图像块大小为  $50 \times 50$ , 其他层图像块大小分别为  $25 \times 25$ 、 $12 \times 12$ 。最终, 结合 3 层金字塔模型, 一幅害虫图像  $Y$  的位置空间金字塔稀疏编码可表示为

$$F_r^l = [\hat{f}_1^l, \hat{f}_2^l, \dots, \hat{f}_p^l], p = 4^{r-1}, r = 2, 3, 4. \quad (9)$$

同理, 害虫图像方向空间金字塔稀疏编码可以通过式 (10) 计算

$$F_r^o = [\hat{f}_1^o, \hat{f}_2^o, \dots, \hat{f}_p^o], p = 4^{r-1}, r = 2, 3, 4. \quad (10)$$

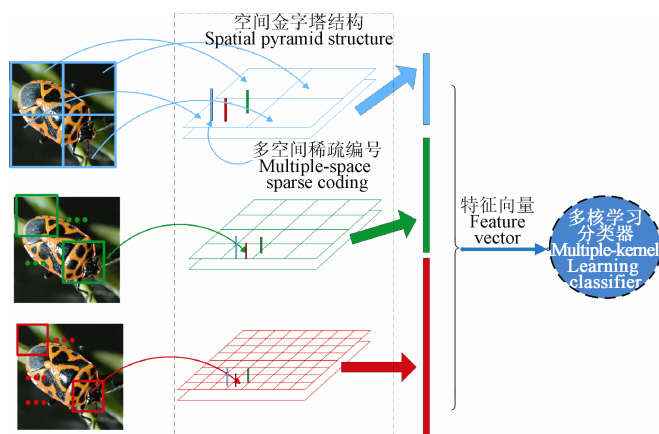


图 2 稀疏编码金字塔结构

Fig.2 Spatial pyramid structure based on sparse coding

### 2.4 基于多核学习的害虫图像识别

在完成害虫图像稀疏编码金字塔特征提取后, 为了更好地融合位置空间与方向空间下的特征, 本文利用多核学习分类方法来实现害虫图像的训练与识别。多核学习是支持向量机 (SVM) 方法的扩展。假设有  $m$  个样本的特征训练数据集  $\{(x_1, y_1), \dots, (x_m, y_m)\}$ ,  $y_i \in \{+1, -1\}$  是标记信息, 则 SVM 二元决策方程定义如下

$$f(x) = W^T R(x) + b. \quad (11)$$

式中  $R(x)$  是样本  $x$  到高维特征空间的映射函数。进一步通过解以下优化问题得到二元分类器

$$\min_{W, b, \xi} \frac{1}{2} \|W\|^2 + C \cdot \sum_{i=1}^m \xi_i \quad \text{s.t.} \quad y_i (W^T R(x) + b) + \xi_i \geq 1, 0 \leq i \leq m, \xi_i \geq 0. \quad (12)$$

式中  $W$  是一个超平面法向量,  $\xi_i$  为松弛变量,  $C$  是正则化参数,  $b$  是偏移量,  $(W, b)$  通过解式 (13) 优化问题获得

$$\max \left\{ \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i,j=1}^m \alpha_i \alpha_j y_i y_j \langle R(x_i), R(x_j) \rangle \right\} \quad \text{s.t.} \quad \sum_{i=1}^m y_i \cdot \alpha_i = 0, 0 \leq i \leq m, 0 \leq \alpha_i \leq C \quad (13)$$

其中,  $O(x_i, x_j) = \langle R(x_i), R(x_j) \rangle$  是核函数, 该文采用如下高斯核函数。

$$O(x_i, x_j) = \exp\left(-\delta^2 \|x_i - x_j\|^2\right). \quad (14)$$

式中  $\delta^2$  是核参数。假设一个测试样本为  $x$ ，则二元分类器表达式如下

$$f(x) = \sum_{i=1}^m \alpha_i \cdot y_i \cdot O(x_i, x) + b. \quad (15)$$

为了融合位置空间与方向空间的稀疏编码特征，通过联合多个核函数方式完成不同特征下权值的自适应学习。假设有  $m$  个害虫图像训练样本，在金字塔模型下，对应的位置空间稀疏编码特征为  $\{\hat{F}_1^l, \hat{F}_2^l, \dots, \hat{F}_m^l\}$  其中  $\hat{F}_i^l = \{F_2^l, F_3^l, F_4^l\}$ ， $F_r^l = [\hat{f}_1^l, \hat{f}_2^l, \dots, \hat{f}_p^l]$ ， $p = 4^{r-1}$ ， $i = 1, 2, \dots, m$ 。

对应的方向空间稀疏编码特征为  $\{\hat{F}_1^o, \hat{F}_2^o, \dots, \hat{F}_m^o\}$  其中  $\hat{F}_i^o = \{F_2^o, F_3^o, F_4^o\}$ ， $F_r^o = [\hat{f}_1^o, \hat{f}_2^o, \dots, \hat{f}_p^o]$ ， $p = 4^{r-1}$ ， $i = 1, 2, \dots, m$ 。则最终对应的多核学习方程如下

$$\min_{\{w_s\}, b, \xi} \frac{1}{2} \sum_s \frac{1}{\omega_s} \|w_s\|^2 + C \cdot \sum_{i=1}^m \xi_i$$

$$\text{s.t. } y_i \left( \sum_s \langle w_s, R_s(\hat{F}_i^s) \rangle \right) + \xi_i \geq 1, 0 \leq i \leq m, \xi_i \geq 0. \quad (16)$$

$\sum_s \omega_s = 1, \omega_s \geq 0, s = \{l, o\}$   
式中  $\omega_s$  对应不同特征模式下基本核函数  $O_s(\hat{F}_i^s, \hat{F}_j^s) = \langle R_s(\hat{F}_i^s), R_s(\hat{F}_j^s) \rangle$  的权值。最终，害虫图像测试样本  $x$  通过式 (17) 多核学习分离器进行分类

$$f(x) = \sum_{i=1}^m \omega_s \cdot \alpha_i \cdot y_i \cdot O_s(H_i, x) + b. \quad (17)$$

本文算法主要步骤如下：

#### 1) 字典学习

①从大量样本图像中随机抽取  $M$  个图像块；②对收集来的每个图像块进行亮度、对比度以及白化等步骤的预处理；③通过式 (4) 应用非监督学习算法 K-SVD 创建完备字典  $D = [d_1, d_2, \dots, d_k] \in \mathbb{R}^{n \times M}$ 。

#### 2) 害虫图像多空间稀疏编码

①对训练样本的每个图像块进行亮度、对比度以及白化等步骤的预处理；②通过式 (5) 对训练样本进行稀疏编码；③根据式 (6) ~ (7) 对训练样本分别进行位置空间稀疏编码以及方向空间稀疏编码；④根据式 (9) ~ (10) 获得最终害虫图像样本的金字塔结构下的稀疏编码特征分布。

#### 3) 多核学习与识别

①将步骤 1 和步骤 2 计算的训练样本编码特征通过式 (16) 训练一个多核学习分离器；②对于测试样本图像  $x$ ，通过计算完备字典的位置空间与方向空间稀疏编码，并经过金字塔模型特征提取后，通过式 (17) 得出最终预测的分类结果。

### 3 试验结果及分析

#### 3.1 害虫图像测试样本及识别系统

本文针对斑须蝽、大稻缘蝽、大青叶蝉、褐飞虱、卷叶螟、大螟、葱蓟马等 35 种农作物害虫，以统一标准获取 8 000 余幅图像，建立样本库，每类害虫样本图像数量在 200 ~ 300 幅之间。试验中，每类选择 150 幅图像作为训练样本，剩余图像作为测试样本。特别地，针对同

一害虫的不同虫龄，由于在形态上的差异性较大的特点，利用该文方法，通过标记同一害虫的不同子类别来处理，即对应同一类别害虫的幼虫与成虫分别标记为该类别的 2 个子类，然后通过该文方法进行识别。测试样本具体信息及识别结果如表 1 所示。

本文待测试害虫图像预处理包括：图像去噪、分割。其中，害虫图像样本分割算法采用 GrabCut 算法<sup>[23]</sup>，该算法是一种交互式图像分割算法，通过用户选定待分割图像的目标区域后，经过计算图像中的纹理、颜色信息和边界反差信息，实现自动分割结果，如图 3 所示。所有的实验室数据，包括害虫图像人工分类标记、标准化和样本图像预处理，均由安徽农业科学院农业经济与信息研究所完成。具体方法为：利用单反数码相机和智能手机，在自然光和闪光灯补光 2 种光线环境下，以植株、土壤、天空等自然环境为背景，在田间直接采集虫害生态图像；再利用图像处理软件对采集到的图像进行元数据标引、裁切、人工分类等预处理。试验中图像大小标准为 200×200，图像的格式为 JPG 格式。试验程序在 Visual C++ 2010 与 Open CV2.4.2 环境下实现。

本文试验的系统界面如图 3 所示，该系统主要实现了对 4 种大田农作物共 35 种害虫的识别。首先，利用数码相机或智能手机，在自然光环境下，直接采集田间虫害生态图像。通过计算机或者手机将待识别害虫图像读入自动识别系统，然后通过人工标记矩形框来选定待分割图像的目标区域，采用 GrabCut 算法进行自动分割，最后使用该文提到的算法对测试样本进行自动识别，给出识别结果与对应的害虫防治方法。

#### 3.2 不同特征下害虫识别准确率

为了验证不同特征下害虫识别方法的有效性，分别提取害虫的颜色直方图<sup>[24]</sup>、LBP 局部纹理<sup>[25]</sup>、形状特征<sup>[26]</sup>以及本文所提的稀疏编码特征来构造训练特征样本，通过 SVM 线性分类器进行训练与测试。表 2 给出了不同特征模式下的 35 种害虫识别精度。从表 1 可知，通过害虫图像纹理特征进行训练学习，识别的准确率较低；通过害虫图像的颜色或形状特征进行训练学习，识别的准确率较高。其主要原因是纹理特征往往反映的是目标物体局部信息，颜色与形状特征通常作为描述目标物体的全局统计特征。显然，对于背景信息较为丰富的害虫图像，全局特征尤为重要，然而局部特征亦可作为全局特征有益的补充。与其他 3 种特征提取方法相比，基于该文所提特征表示方法的识别精度均提高了 9% 以上，这主要得益于多空间稀疏编码在害虫图像特征表示上的优势。

#### 3.3 不同方法下 35 种害虫的识别准确率

为了验证该文所提的害虫识别方法的有效性，在同一数据集上，分别采用本文方法 (Ours)、支持向量机方法<sup>[6]</sup> (SVM)、神经网络方法<sup>[27]</sup> (NN)、稀疏编码方法<sup>[17]</sup> (SC) 进行害虫识别准确率比较。表 1 给出了 4 种方法下的 35 种害虫识别准确率。由表 1 不难发现，神经网络方法以及支持向量机方法的识别准确率不高。文献[17]所提的稀疏表示方法与本文方法较为接近，但仍然低于本文方法的平均识别准确率。

表 1 35 种害虫测试样本细节、不同特征和方法下的害虫识别准确率  
Table 1 Details of 35 analyzed insect species and pest recognition rate under different features and methods

害虫样本 Species ID	害虫名称 Species names	虫态 Stages	测试样 本数量 Number of test samples	不同特征下的害虫识别准确率 Pest recognition rate under different features/%				不同方法下的害虫识别准确率 Pest recognition rate under different methods/%			
				颜色 Color	纹理 Texture	形状 Shape	本文方法 Ours	稀疏编码 方法 SC	神经网络 方法 NN	支持向量 机方法 SVM	本文方法 Ours
1	斑须蝽 <i>Dolycoris baccarum</i> (Linnaeus)	成虫 Adult	53	66.0	58.4	62.2	75.5	68.0	73.4	74.2	75.5
2	斑衣蜡蝉 <i>Lycorma delicatula</i> (White)	成虫 Adult	135	67.1	61.5	60.1	74.1	69.1	72.5	73.3	74.1
3	菠萝粉蚧 <i>Dysmicoccus brevipes</i> (Cockerell)	成虫 Adult	61	69.1	53.1	56.3	78.7	70.1	73.1	75.3	78.7
4	菜蝽 <i>Eurydema dominulus</i> (Scopoli)	成虫 Adult	125	64.0	57.6	53.6	72.0	68.0	70.6	71.6	72.0
5	菜粉蝶 <i>Pieris rapae</i> (Linnaeus)	幼虫 Larva	89	76.7	69.7	65.7	90.1	77.2	78.4	80.3	90.1
6	蚕豆象 <i>Bruchus rufimanus</i> Boheman	成虫 Adult	60	66.3	61.2	56.1	83.3	69.3	74.2	76.1	83.3
7	茶翅蝽 <i>Halymorpha halys</i> (Stål)	成虫 Adult	100	58.0	69.0	62.0	82.0	70.6	75.0	81.2	82.0
8	尘污灯蛾 <i>Spilosoma obliqua</i> (Walker)	成虫 Adult	78	63.9	61.3	60.0	76.9	65.9	69.3	75.6	76.9
9	赤条蝽 <i>Graphosoma rubrolineata</i> (Westwood)	成虫 Adult	90	58.0	71.2	69.0	80.0	71.7	73.2	78.0	80.0
10	臭椿皮蛾 <i>Eligma narcissus</i> (Cramer)	成虫 Adult	65	60.9	63.9	56.4	76.9	67.9	68.9	75.4	76.9
11	葱黄寡毛跳甲 <i>Luperomorpha suturalis</i> Chen	成虫 Adult	135	67.4	59.7	52.7	66.7	63.4	64.7	65.4	66.7
12	葱蓟马 <i>Thrips tabaci</i> Lindeman	成虫 Adult	121	68.0	58.4	66.4	74.4	69.0	70.4	71.8	74.4
13	大稻缘蝽 <i>Leptocorisa acuta</i> (Thunberg)	成虫 Adult	152	63.0	58.8	63.7	65.8	63.1	63.8	64.3	65.8
14	大螟 <i>Sesamia inferens</i> (Walker)	幼虫 Larva	69	65.4	61.7	67.8	71.2	68.5	69.4	70.4	71.2
15	大青叶蝉 <i>Cicadella viridis</i> (Linnaeus)	成虫 Adult	82	69.8	63.8	75.6	85.4	72.8	75.8	77.6	85.4
16	稻赤斑沫蝉 <i>Callitettix versicolor</i> (Fabricius)	成虫 Adult	79	57.7	62.2	67.4	75.9	60.7	62.7	73.4	75.9
17	稻黑蝽 <i>Scotinophara lurida</i> (Burmeister)	成虫 Adult	55	65.6	60.2	78.2	81.8	68.6	77.2	78.9	81.8
18	稻棘缘蝽 <i>Cletus punctiger</i> (Dallas)	成虫 Adult	64	68.5	60.5	59.8	78.1	69.5	70.5	76.8	78.1
19	稻绿蝽 <i>Nezara viridula</i> (Linnaeus)	成虫 Adult	79	56.4	68.1	61.6	75.9	66.4	69.1	71.6	75.9
20	稻铁甲 <i>Dicladispa armigera</i> (Olivier)	成虫 Adult	127	65.2	60.4	70.0	77.2	67.2	69.4	70.8	77.2
21	稻纵卷叶螟 <i>Cnaphalocrocis medinalis</i> (Guenée)	幼虫 Larva	124	71.0	65.2	61.2	80.6	75.0	78.2	79.8	80.6
22	点蜂缘蝽 <i>Riptortus pedestris</i> (Fabricius)	成虫 Adult	126	72.0	68.0	70.6	79.4	74.0	78.0	78.6	79.4
23	豆荚野螟 <i>Maruca testulalis</i> Gryer	幼虫 Larva	92	71.3	67.2	68.2	74.6	69.2	71.8	73.2	74.6
24	豆突眼长蝽 <i>Chauliops fallax</i> Scott	成虫 Adult	66	65.3	59.3	69.8	75.8	68.3	72.3	74.6	75.8
25	短额负蝗 <i>Atractomorpha sinensis</i> Bolvar	成虫 Adult	75	74.8	63.1	72.2	80.0	76.8	79.1	73.2	80.0
26	二化螟 <i>Chilo suppressalis</i> (Walker)	幼虫 Larva	87	72.8	59.3	69.1	76.2	72.3	70.2	71.7	76.2
27	扶桑绵粉蚧 <i>Phenacoccus solenopsis</i> Tinsley	成虫 Adult	67	62.9	58.4	68.9	74.6	66.9	72.4	73.9	74.6
28	甘薯天蛾 <i>Agrius convolvuli</i> (Linnaeus)	成虫 Adult	59	72.9	64.4	74.6	81.4	76.9	77.4	79.6	81.4
29	瓜绢螟 <i>Diaphania indica</i> (Saunders)	成虫 Adult	135	68.5	58.0	65.0	74.1	69.5	72.0	73.2	74.1
30	广二星蝽 <i>Stollia ventralis</i> (Westwood)	成虫 Adult	127	64.2	55.7	64.2	70.6	65.7	66.4	68.6	70.6
31	禾谷缢管蚜 <i>Rhopalosiphum padi</i> (Linnaeus)	成虫 Adult	54	71.9	62.4	77.6	83.3	75.9	78.4	80.6	83.3
32	褐飞虱 <i>Nilaparvata lugens</i> (Stål)	成虫 Adult	126	69.0	59.4	70.6	77.8	70.0	74.4	76.6	77.8
33	红袖蜡蝉 <i>Diostrombus politus</i> Uhler	成虫 Adult	78	65.2	70.4	58.7	79.5	68.2	70.7	73.7	79.5
34	豆蚜 <i>Aphis craccivora</i> Koch	成虫 Adult	125	59.2	64.4	57.6	71.2	63.2	64.4	67.6	71.2
35	黄曲条跳甲 <i>Phyllotreta striolata</i> (Fabricius)	成虫 Adult	50	70.0	74.0	62.0	80.0	72.0	74.0	79.2	80.0
平均识别率				67.1	62.8	65.2	78.0	70.0	72.7	75.4	78.2



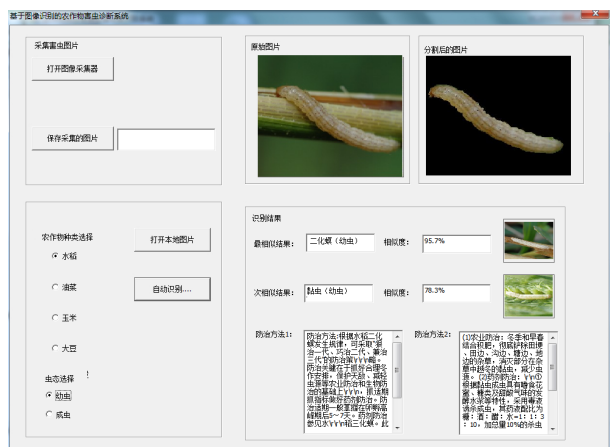


图 3 系统界面

Fig.3 GUI of system

### 3.4 其他数据集上识别的准确率

为了进一步验证所提算法的有效性, 该文对比了 4 种常见算法在 2 个不同数据集上的平均识别准确率。文献[28]提供的数据集包含 221 种昆虫, 这些昆虫图像样本均在实验室环境下采集, 经图像分割等预处理。因此, 样本图像的背景较为单一, 识别也相对容易。文献[29]提供的数据集包含 20 种蝴蝶样本, 由于这些害虫图像外观形状极其相似, 因此通过单一提取形状特征很难进行识别。由表 2 可知, 本文所提算法在 2 个数据集上均获得了更高的平均识别准确率, 该文所提方法可使平均识别精度分别提高 13.4%、9.5%以及 14.1%。

表 2 不同数据库下平均识别准确率

Table 2 Pest recognition rate in different datasets

数据集 Datasets	支持向量机 SVM	神经网络 NN	稀疏编码 SC	本文方法 Our method
文献[28] Reference[28]	70.2	78.7	82.9	88.3
文献[29] Reference[29]	65.5	72.9	80.3	90.0
平均识别率	67.9	75.8	81.6	89.2

## 4 结 论

针对传统特征提取方法在害虫图像识别中所存在的不足, 本文结合图像空间金字塔模型以及稀疏编码思想, 提出了一种新的能够有效融合多空间稀疏编码的农作物害虫图像识别方法。该方法通过非监督特征学习构建过完备字典, 并结合图像金字塔模型, 计算害虫图像的位置空间与方向空间上的稀疏编码特征, 分别构造相应的特征训练样本矩阵, 最后结合多核学习框架, 实现了不同空间稀疏编码特征的有效融合。在不同特征下、相同数据集以及不同数据集上的 3 类试验结果表明, 与其他方法相比, 该文所提方法可使平均识别精度分别提高 13.4%、9.5%以及 14.1%。

### 参 考 文 献

- [1] 雷仲仁, 郭予元, 李世访. 中国主要农作物有害生物名录 [M]. 北京: 中国农业科学技术出版社, 2014.
- [2] 赵汗青, 沈佐锐, 于新文. 数学形态学在昆虫分类学上的应用研究 I: 在目级阶元上的应用研究[J]. 昆虫学报, 2003, 46(1): 45—50.
- [3] 赵汗青, 沈佐锐, 于新文. 数学形态学在昆虫分类学上的应用研究 II: 在总科级阶元上的应用研究[J]. 昆虫学报, 2003, 46(2): 201—208.
- [4] 沈佐锐, 赵汗青, 于新文. 数学形态学在昆虫分类学上的应用研究 III: 在科阶元上的应用研究[J]. 昆虫学报, 2003, 46(3): 339—344.
- [5] 杨红珍, 张建伟, 李湘涛, 等. 基于图像的昆虫远程自动识别系统的研究[J]. 农业工程学报, 2008, 24(1): 188—192.
- [6] 胡玉霞, 张红涛. 基于模拟退火算法-支持向量机的储粮害虫识别分类[J]. 农业机械学报, 2008, 39(9): 108—111.
- [7] 毛文华, 郑永军, 张银桥, 等. 基于机器视觉的草地蝗虫识别方法[J]. 农业工程学报, 2008, 24(11): 155—158.
- [8] 张红涛, 毛罕平, 邱道尹. 储粮害虫图像识别中的特征提取[J]. 农业工程学报, 2009, 25(2): 126—130.
- [9] Zhao Hanqing, Shen Zuorui, Yu Xinwen. Use of math-morphological features in insect taxonomy I: At the order level[J]. Acta Entomologica Sinica, 2003, 46(1): 45—50. (in Chinese with English abstract)
- [10] Zhao Hanqing, Shen Zuorui, Yu Xinwen. Use of math-morphological features in insect taxonomy II: At superfamily level[J]. Acta Entomologica Sinica, 2003, 46(2): 201—208. (in Chinese with English abstract)
- [11] Shen Zuorui, Zhao Hanqing, Yu Xinwen. Use of math-morphological features in insect taxonomy III: At the family level[J]. Acta Entomologica Sinica, 2003, 46(3): 339—344. (in Chinese with English abstract)
- [12] Yang Hongzhen, Zhang Jianwei, Li Xiangtao, et al. Remote automatic identification system based on insect image[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2008, 24(1): 188—192. (in Chinese with English abstract)
- [13] Hu Yuxia, Zhang Hongtao. Recognition of the stored-grain pests based on simulated annealing algorithm and support vector machine[J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(9): 108—111. (in Chinese with English abstract)
- [14] Mao Wenhua, Zheng Yongjun, Zhang Yinqiao, et al. Grasshopper detection method based on machine vision[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2008, 24(11): 155—158. (in Chinese with English abstract)

- Zhang Hongtao, Mao Hanping, Qiu Daoyin. Feature extraction for the stored-grain insect detection system based on image recognition technology[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2009, 25(2): 126—130. (in Chinese with English abstract)
- [9] 竺乐庆, 张真, 张培毅. 基于颜色直方图及双树复小波变换 (DTCWT) 的昆虫图像识别[J]. 昆虫学报, 2010, 3(1): 91—97.
- Zhu Leqing, Zhang Zhen, Zhang Peiyi. Image identification of insects based on color histogram and dual tree complex wavelet transform (DTCWT)[J]. Acta Entomologica Sinica, 2010, 3(1): 91—97. (in Chinese with English abstract)
- [10] Zhu Leqing, Zhang Zhen. Insect recognition based on integrated region matching and dual tree complex wavelet transform[J]. Journal of Zhejiang University: Science C, 2011, 12(1): 44—53.
- [11] Jiangning Wang, Congtian Lina, Liqiang Jia, et al. A new automatic identification system of insect images at the order level[J]. Knowledge-Based Systems, 2012, 33(1): 102—110.
- [12] Chenglu Wen, Daniel Guyer. Image-based orchard insect automated identification and classification method[J]. Computers and Electronics in Agriculture, 2012, 89(1): 110—115.
- [13] Tao Liu, Wen Chen, Wei Wu, et al. Detection of aphids in wheat fields using a computer vision technique[J]. Biosystems Engineering, 2016, 141(1): 82—93.
- [14] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210—227.
- [15] 韩安太, 郭小华, 廖忠, 等. 基于压缩感知理论的农业害虫分类方法[J]. 农业工程学报, 2011, 27(6): 203—207.
- Han Antai, Guo Xiaohua, Liao Zhong, et al. Classification of agricultural pests based on compressed sensing theory[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2011, 37(6): 203—207. (in Chinese with English abstract)
- [16] Han Antai, Peng Hui, Li Jianfeng, et al. Recognition of pests based on compressive sensing theory[C]//2011 IEEE 3rd International Conference on Communication Software and Networks (ICCSN), 2011: 263—266.
- [17] 胡永强, 宋良图, 张洁, 等. 基于稀疏表示的多特征融合害虫图像识别[J]. 模式识别与人工智能, 2014, 27(11): 985—992.
- Hu Yongqiang, Song Liangtu, Zhang Jie, et al. Pest image recognition of multi-feature fusion based on sparse representation[J]. Pattern Recognition and Artificial Intelligence, 2014, 27(11): 985—992. (in Chinese with English abstract)
- [18] Xie Chengjun, Zhang Jie, Li Rui, et al. Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning[J]. Computers and Electronics in Agriculture, 2015, 119(1): 123—132.
- [19] Elhamifar E, Vidal R. Robust classification using structured sparse representation[C]//Computer Vision and Pattern Recognition (CVPR), 2011: 1873—1879.
- [20] Adam Coates, Andrew Ng, Honglak Lee. An analysis of single-layer networks in unsupervised feature learning[C]//International Conference on Artificial Intelligence and Statistics, 2011: 215—223.
- [21] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311—4322.
- [22] Adam Coates, Andrew Ng. The importance of encoding versus training with sparse coding and vector quantization[C]//International Conference on Machine Learning, 2011: 921—928.
- [23] Carsten R, Vladimir K, Andrew B. “Grabcut”—interactive foreground extraction using iterated graph cuts[J]. Acm Transactions on Graphics, 2004, 23(3): 307—312.
- [24] Swain M J, Ballard D H. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1): 11—32.
- [25] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971—987.
- [26] Chan Yuk-ming. Shape-based Image Retrieval in Iconic Image Databases[M]. Hongkong: The Chinese University of Hong Kong, 1999.
- [27] 邱道尹, 张成花, 张红涛, 等. 神经网络在储粮害虫识别中的应用[J]. 农业工程学报, 2003, 19(1): 142—144.
- Qiu Daoyin, Zhang Chenghua, Zhang Hongtao, et al. Application of neural networks in the recognition of stored grain pests[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2003, 19(1): 142—144. (in Chinese with English abstract)
- [28] Wang Jiangning, Lin Congtian, Ji Liqiang, et al. A new automatic identification system of insect images at the order level[J]. Knowledge-Based Systems, 2012, 33(1): 102—110.
- [29] Xiao Bin, Ma Jianfeng, Cui Jiangtao. Combined blur, translation, scale and rotation invariant image recognition by Radon and pseudo-Fourier-Mellin transforms[J]. Pattern Recognition 2012, 45(1): 314—321.

## Recognition for insects via spatial pyramid model using sparse coding

Xie Chengjun<sup>1</sup>, Li Rui<sup>1\*</sup>, Dong Wei<sup>2</sup>, Song Liangtu<sup>1</sup>, Zhang Jie<sup>1</sup>, Chen Hongbo<sup>1</sup>, Chen Tianjiao<sup>1</sup>

(1. Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China;

2. Agricultural Economy and Information Research Institute, Anhui Academy of Agricultural Sciences, Hefei 230031, China)

**Abstract:** Automatic classification of insect species in field crops such as corn, soybean, wheat, and canola is more difficult than the generic object classification because of complex background in field and high appearance similarity among insect species. In this paper, we propose an insect recognition system on the basis of advanced sparse coding and spatial pyramid model. We firstly learn features from a large amount of unlabeled insect image patches to construct an over-complete dictionary. The sparse coding of insect image patches is obtained by encoding over the dictionary. To enhance discriminative ability of the sparse coding, we then apply multiple scales of filters coupled with different spaces. Finally, the multiple space features of sparse coding are seamlessly embedded into a multi-kernel framework for robust classification. Traditionally, insect recognition has mainly relied on manual identification by expert entomologists. However, for laymen without a thorough understanding of the terminology of insect taxonomy and morphological characteristics, it is hard to discriminate insect categories at the species level. Therefore, effective identification of insects is a key issue that needs to be well addressed. To improve the recognition accuracy, we develop an insect recognition system using advanced sparse coding, spatial pyramid model and multiple-kernel learning techniques. Different from traditional feature representation, a novel feature representation that is multiple-space sparse coding of insect objects is proposed by this work. The work flow of our method can be decomposed into 2 stages. The first stage focuses on image or insect object representation. At this stage, the features of insect images are extracted using advanced sparse coding and spatial pyramid model. The second stage, which deals with effective fusion of multiple insect-categorization features, constructs a kernel-level fusion classifier using all the sparse coding features. At the first stage, for an insect image given, the features of insect images are extracted firstly. The features are then represented as a linear combination of the corresponding training feature dictionary. Then, a multiple-space sparse coding with spatial pyramid model is used to represent insect image in a joint sparse way over all the features. In this process, for the object image, an over-complete dictionary with unlabeled insect images is learned first. Then, the local image patches of the insect object are represented by their sparse codes with the training dictionary. Despite the fact that appearance is modeled using local patches, the global structure information is necessary for accurate insect identification. Consequently, insect appearance is represented by concatenating the location and orientation sparse-coding of all image patches. To obtain the more compact representations of insect images, we use spatial pyramid model at multi-scale levels, which achieves better robustness to noise and clutter, and thus better copes with severe variations in the pose, scale or rotation. In this paper, we use the 3-scale level pyramid to represent insect image. At the top level of the pyramid, there are 4 image patches which represent the whole image. Each image block size is 50×50. The middle level contains 16 equal size non-overlapping image patches, for which each image block size is 25×25; and the bottom level has 64 image patches for which each image block size is 12×12. Then, the local image features of the 3-level pyramid are combined to represent the insect appearance. The larger scale level provides better geometric features when the classifying insects undergo large appearance variations, while smaller scale level obtains finer features. Finally, the features from fine to coarse levels across different scales are concatenated together to generate the final feature representation of insect images for insect classification. At the second stage, the multiple-kernel learning approach is adopted to combine multiple-space sparse coding. As different features of insect images contribute differently to the classification of insect species, the multiple-space sparse coding technique can combine multiple features of insect species to enhance the recognition performance. Given positive and negative insect samples, the features are extracted. Local image patches of the samples are then represented by multiple-space sparse coding using the corresponding training dictionary. Finally, the multiple-kernel learning classifier is constructed by learning the multiple-space sparse coding of the negative and positive samples for insect categorization and recognition. To meet the need of practical insect image identification, we collected insect images covering various species across several common field crops including corn, soybean, wheat, and canola. Samples of 35 common pest species found in field crops were collected, such as *Pieris rapae* (Linnaeus), and *Leptocorisa acuta* (Thunberg). Experimental results showed that our proposed method performed well on the classification of insect species, and outperformed the state-of-the-art methods of the generic insect recognition. Our method improved the recognition rate by more than 9% compared to other methods for the same data sets. In addition, the proposed method had also a good performance and enhanced the average recognition accuracy by 14.1% for the different data sets.

**Keywords:** image recognition; algorithms; pest control; dictionary learning; sparse coding; spatial pyramid model