

# 基于多特征融合的玉米前期图像的旱情识别

路志英, 刘书辰, 官志宏

天津大学电气自动化与信息工程学院天津市过程检测与控制重点实验室, 天津 300072

**摘要** 为实现对玉米植株旱情的分析, 针对目前农业干旱指标涉及领域较为广泛、获取困难的研究现状, 提出了一种基于多特征融合的玉米前期图像旱情识别方法。以正常和特旱两种情况的玉米植株图像为样本, 采用经典 K-means 算法对玉米植株图像提取感兴趣区域; 进而提取分割后的玉米植株图像, 包括颜色、奇异值分解(SVD)、纹理等共计 20 维特征; 采用遗传算法对 20 维特征选择有效特征子集; 最后针对有效特征子集建立了基于最小二乘支持向量机的判别模型, 获取了玉米植株图像的旱情信息。将单个特征(颜色、SVD、纹理)直接融合之后的特征以及利用主成分分析法的特征选择作为对比实验, 平均识别正确率分别为 0.9503、0.9627、0.9771、0.9460、0.9745, 而采用遗传算法进行特征选择后, 最终寻到最优解为 9 维特征, 平均识别正确率为 0.9903。结果表明, 运用图像处理技术可以对旱情进行识别, 取得了较好的效果, 为农业旱情的识别提供了新思路。

**关键词** 图像处理; 旱情识别; 多特征融合; 遗传算法

**中图分类号** TP3-05 **文献标识码** A

**doi:** 10.3788/LOP54.081005

## Drought Identification Based on Multi-Features Fusion for Early Maize Images

Lu Zhiying, Liu Shuchen, Gong Zhihong

*Tianjin Key Laboratory of Process Measurement and Control, School of Electrical and Information Engineering,  
Tianjin University, Tianjin 300072, China*

**Abstract** In order to analyze the drought of maize plants and aiming at the difficulty and broad in recognizing agricultural drought, we propose a method to identify the drought of maize plants based on the multi-features fusion. The images of normal and seriously drought plants are taken as samples. The K-means algorithm is used to extract the interesting areas of maize plant images. And, the features of the pictures are extracted after image segmentation, including colors, singular value decomposition (SVD) and textures, a total of 20 dimensional features. The genetic algorithm is used to select a effective features subset of 20 dimensional features. Finally, the discrimination model based on least squares support vector machine is established for the effective features subset and images of maize plant drought are obtained. The single feature (color, SVD, texture) after directly fusion and using principal component analysis for feature selection are performed as comparative experiments, the average recognition accuracies are 0.9503, 0.9627, 0.9771, 0.9460, 0.9745, respectively. The genetic algorithm is used to select the features, and finally finds 9 dimensional features as the optimal solution. The average recognition accuracy is 0.9903. The result shows that this image processing technology can identify the drought situation of the maize plants effectively and efficiently. And it also provides a new idea for the drought identification of maize plants.

**Key words** image processing; drought identification; multi-feature fusion; genetic algorithm

**OCIS codes** 100.2960; 100.2000; 150.6044

## 1 引言

目前环境问题迫在眉睫, 其中频发的重大干旱灾难尤为突出<sup>[1]</sup>。干旱本身作为一种复杂的现象难以直接观测其发生时间、发展过程以及影响范围。国内外对旱情的监测手段, 从旱情统计上报到测定土壤含水量, 发展到如今的遥感图像技术, 利用多时相、高光谱遥感数据反映大面积的图像信息<sup>[2]</sup>, 实现旱情

**收稿日期:** 2017-02-24; **收到修改稿日期:** 2017-03-29

**作者简介:** 路志英(1964—), 女, 博士, 教授, 主要从事数据挖掘、图像处理方面的研究。E-mail: luzhy@tju.edu.cn

分析。

但是目前遥感图像仍然存在成本高、处理复杂、局部分析困难等问题。本文利用室外摄像机获取图像,虽然图像相对分辨率较低,但仍包含作物在不同程度的干旱胁迫状况下,形态和颜色随之发生改变的图像信息。通过数字图像处理技术,捕获作物本身的状态,进而反映作物的受旱程度。数字图像处理技术是为了达到某种预想的目的而对图像进行运算和处理的一种过程,主要应用在刑侦技术、通讯技术、医学影像、生物识别、气象等方面。在农业中的应用主要在农作物状态监测、作物形状识别分析、作物病虫害监测等领域。例如 Lu 等<sup>[3]</sup>在玉米阶段检测、产量估计、机械去雄以及玉米雄穗表型参数等方面取得较大的研究成果; Camargo 等<sup>[4]</sup>通过对棉花病虫害叶片图像的一系列处理,采用支持向量机的方法实现了棉花病害的分类,取得了较好的效果。

## 2 玉米植株图像获取与图像预处理

### 2.1 图像采集

采集自然光照下正常和特旱两种情况的玉米植株图像作为样本(图 1),提取对其感兴趣的区域(叶片),利用颜色直方图、奇异值和灰度共生矩阵对分割之后的玉米叶片图像提取特征,结合遗传算法(GA)进行多特征融合构建最小二乘支持向量机(LSSVM)分类器,实现了特征子集的优选以及玉米植株图像的旱情识别。



图 1 (a)(b)正常玉米植株;(c)(d)特旱玉米植株

Fig. 1 (a)(b) Normal maize plants; (c)(d) seriously drought maize plants

图 1 中所提供的玉米植株样本图像为不同距离下的拍摄结果并且部分存在遮挡物。因此通过图像对玉米植株进行旱情判别的方法,对提取图像特征及构建分类器提出了较高的要求。

所获取的样本为玉米品种郑单 958,为保证玉米植株正常出苗,在其生长前期土壤水分正常供应。研究对象为生育前期(出苗-大喇叭口时期),生育阶段分为两个土壤水分处理水平,分别为正常植株以及特旱植株。正常玉米植株土壤水分控制范围为 65%~80%(即土壤水分含量占田间持水量的百分数),特旱玉米植株土壤水分控制为小于 30%。处理期间,根据土壤水分传感器每 5 min 测定土壤体积含水量数据,判断是否需要灌水,当实际测量土壤含水量低于控制下限时,灌水至控制上限。灌水时在水管头上安接一个水表,严格按照计算的灌溉量来灌水。另外,灌水时根据水的流速匀速移动喷水头,保证整个小区灌水均匀。当降雨天气发生时,及时关闭防雨棚,从而实现完全人为控制土壤水分的补充。

图像采集前端选用型号为 WV-SW396AH 的松下 720p 室外网络快球摄像机端,实验图片的采集每 2 h 进行一次,通过预先设定好的角度采集每一池玉米的生长图片,从早上 6:00 开始到傍晚 18:00 结束。

### 2.2 叶片提取

利用玉米叶片和土壤背景的显著颜色差异,在 RGB 空间中采用 K-means 算法<sup>[5]</sup>对玉米植株图像进行聚类处理,从而实现玉米叶片的提取。K-means 算法步骤如下:

- 1) 随机选择两个点作为初始质心。
- 2) 将图像中的每个点指派到最近的质心,形成两个簇,重新计算每个簇的质心,所选取的距离量度为欧氏距离,簇的质心取其均值。
- 3) 簇不发生变化或达到最大迭代次数则停止搜索。

K-means 算法计算时间短、速度快、图像分割效果佳。图 2 为预处理前、后的效果图。

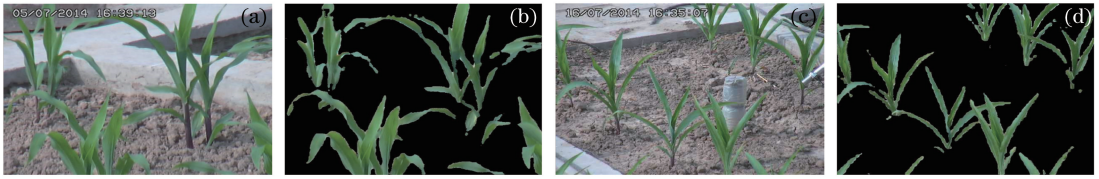


图2 (a)(b)正常植株分割效果图;(c)(d)特旱植株分割效果图

Fig. 2 (a)(b) Normal plant image segmentation effect diagrams; (c)(d) seriously droughted plant image segmentation effect diagrams

### 3 特征提取

#### 3.1 颜色特征提取

正常植株和特旱植株在图像颜色上有显著不同,正常叶片水分供应充足,通常呈嫩绿色;特旱叶片水分供应严重不足,通常呈深绿色,故将提取其颜色特征可以区分正常植株和特旱植株。颜色特征的提取方法主要有颜色直方图、颜色集、颜色矩等。文献[6]中借助图像的直方图实现了特征提取,提取了均值、方差等信息来描述直方图的特征。

本文针对叶片图像的RGB直方图,采用统计学方法,分别统计R直方图、G直方图、B直方图的均值和偏度<sup>[7-8]</sup>信息作为不同类别间的特征差异。

直方图的均值为图像RGB各通道的像素均值,计算公式为

$$m = \sum_{i=0}^{255} i \cdot p_i, \quad (1)$$

式中  $m$  为直方图均值,  $i$  为像素值,  $p_i$  为像素值为  $i$  的个数与图像像素总个数的比值。

直方图的偏度反映统计数据序列的分布规律,具体包括其数据的对称性及偏斜程度。偏度系数以样本的三阶标准化矩来定义,偏度数值的大小反映其偏斜的程度,偏度数值的正负反映其偏斜的方向,计算公式为

$$S = \frac{\sum (X_i - \mu)^3}{N \cdot \sigma^3}, \quad (2)$$

式中  $S$  为直方图偏度,  $X_i$  为样本数据,  $\mu$  为样本数据的均值,  $\sigma$  为样本数据的方差,  $N$  为样本数据总个数。

根据上述计算公式统计R、G、B直方图的均值和偏度共计6维颜色特征。为了更好地观察每个颜色特征的分类效果,从样本点的6个颜色特征中任意选取B均值和G均值特征,以二维图的方式呈现出来,如图3所示。

将图3中二维平面上的点分别投到对应的横纵轴上,则可以看到单一特征样本点的分类效果,从图中可以看出对于每一个单一特征,两类样本之间存在较大的重叠,可分性较差,而从一维特征提升到二维特征则大大降低了两类样本的重叠区域,增强了分类效果。若将B、R、G均值这三维特征映射到三维图中,则如图4所示。

显然三维分类效果相比二维分类效果进一步增强,因此,不同特征间具有一定的互补性,能够较全面地描述图像。

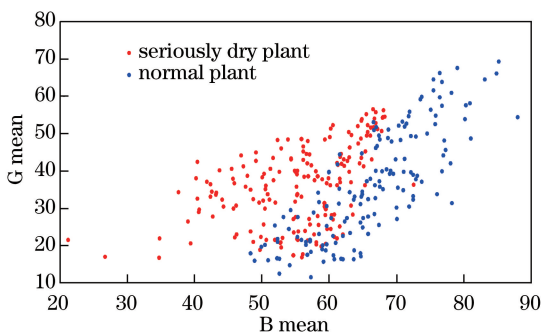


图3 二维特征分类效果图

Fig. 3 Effect of two-dimensional feature classification

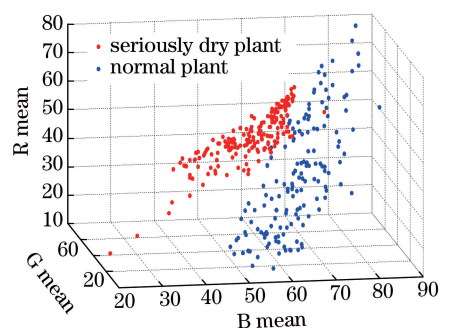


图4 三维特征分类效果图

Fig. 4 Effect of three-dimensional feature classification

### 3.2 纹理特征提取

纹理反映物体表面具有缓慢变化或者周期性变化的表面结构组织排列的属性。在图像分析中,使用纹理表示图像的均匀、细致、粗糙等现象。纹理不同于颜色特征,它能统计中心像素与其周围空间邻域像素的灰度分布规律。纹理特征是用来描述这种空间内灰度变化规律的数字特征。

通过分析不同玉米植株图像,发现正常植株叶片舒展、纹理较为平滑;特旱植株叶片卷曲、纹理相对粗糙。采用统计型纹理中的灰度共生矩阵,通过计算图像中一定距离一定方向的两点灰度之间的相关性,来反映图像的纹理信息。

提取玉米整幅图像的灰度共生矩阵步骤如下:

1) 对分割后玉米植株的图像进行灰度化处理,考虑计算复杂度,对原始灰度图像进行灰度级量化处理,量化为 0~7 共 8 级。

2) 定义一个方向和一个固定的步长,则灰度共生矩阵  $\mathbf{P}(N \times N)$  的第  $(i, j)$  元素值为灰度级  $i$  和  $j$  在灰度图像矩阵内,沿该方向相距该指定步长的两个像素出现的次数总和。

实验选取  $0^\circ$ 、 $45^\circ$ 、 $90^\circ$ 、 $135^\circ$  这 4 个方向,所取步长为 1。由步骤 1) 可知,量化后灰度级为 8,因此生成的灰度共生矩阵为  $8 \times 8$ ,分别得到 4 个方向的灰度共生矩阵。

$0^\circ$  方向上,灰度级为 256 的正常和特旱植株图像灰度共生矩阵所产生的纹理图像如图 5 所示。从图 5 可以看出,正常和特旱植株纹理图像差异较为明显。

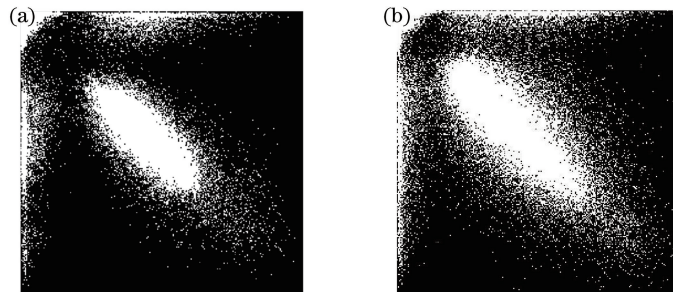


图 5 (a)正常植株纹理图;(b)特旱植株纹理图

Fig. 5 (a) Texture map of normal plant; (b) texture map of seriously drought plant

Haralick 等<sup>[9]</sup>利用灰度共生矩阵提取了图像的 14 种纹理特征。考虑计算复杂度以及对玉米图像旱情识别的分类效果,选取以下 4 种常用特征:对比度、相关性、熵、逆差矩。文献[10]详细给出了这几种纹理特征的定义,正常和特旱特征样本数据对比如表 1 所示。

表 1 正常和特旱样本纹理特征对比表

Table 1 Comparison of texture features of normal and seriously drought samples

	Contrast	Correlation	Entropy	Homogeneity
Normal sample	0.2072	0.9172	0.7558	0.9767
Seriously drought sample	0.6487	0.8768	0.5488	0.7229

从表 1 可以看出,正常和特旱样本的 4 种纹理特征之间存在差异。

### 3.3 奇异值分解

在图像处理领域,奇异值分解(SVD)应用十分广泛<sup>[11]</sup>。SVD 和颜色、纹理特征的目的,就是提取出一个图像最重要的特征<sup>[12]</sup>,而 SVD 是单纯的从矩阵角度考虑。

SVD 公式为

$$\mathbf{A}_{m \times n} = \mathbf{U}_{n \times n} \times \mathbf{\Sigma}_{m \times n} \times \mathbf{V}_{n \times n}^T, \quad (3)$$

式中  $\mathbf{A}$  为  $m \times n$  的矩阵, $\mathbf{U}$  是一个  $m \times m$  的左奇异酉矩阵, $\mathbf{V}^T$  是一个  $n \times n$  的右奇异酉矩阵, $\mathbf{\Sigma}$  为奇异值矩阵,是半正定  $m \times n$  阶对角矩阵。 $\mathbf{\Sigma}$  矩阵对角线上的元素即为矩阵的奇异值,其数值按从大到小顺序排列,奇异值越大对图像的作用越大。

奇异值的衰减速度特别快,绝大多数情况下,前 10% 甚至 1% 的奇异值之和占全部奇异值之和的 99%



以上<sup>[13]</sup>。即,通过前  $r$  个奇异值即可近似描述  $\Sigma$  矩阵 ( $r \ll n$ ),因此,该文选取前  $r$  个奇异值归一化处理作为图像的奇异值特征。经过近似处理的部分 SVD 描述为

$$\mathbf{A}_{m \times n} \approx \mathbf{U}_{m \times r} \Sigma_{r \times r} \mathbf{V}_{r \times n}^T, \quad (4)$$

归一化公式为

$$\hat{\lambda}_i = \frac{\lambda_i}{\lambda_{\max}} = \frac{\lambda_i}{\lambda_1}, i = 1, 2, \dots, r, \quad (5)$$

式中  $\lambda_{\max}$  为图像的最大奇异值(即奇异值矩阵中的第一个元素),  $\lambda_i$  为第  $i$  个奇异值。

上述描述中涉及到一个重要参数,即图像奇异值的特征数目  $r$ 。所获取到的玉米植株图像大小均为  $640 \times 480$ ,因此得到的奇异值的个数为 480。 $r$  通过分类识别成功率进行确定,即每次实验选用前  $n$  个奇异值作为特征来进行训练 ( $n = 1, 2, \dots, 30$ ),将得到的 30 个识别成功率进行比较,选择出较为合适的前  $n$  个奇异值将其归一化作为玉米植株图像的特征值。图 6 给出了识别正确率和  $r$  取值的关系曲线。

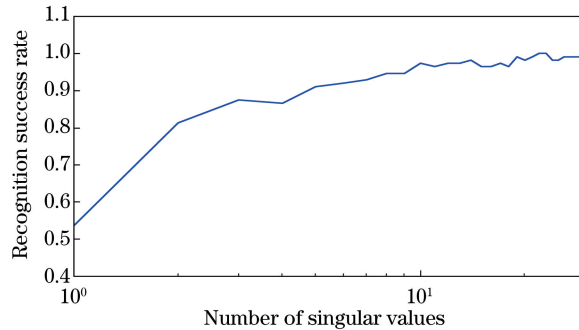


图 6 识别成功率与奇异值个数  $r$  取值的关系曲线

Fig. 6 Relationship curve between recognition success rate and the number of singular values

从图 6 可以看出,奇异值个数大于 10 后识别正确率能达到 90% 以上,考虑特征维数过大有可能增加计算复杂度,兼顾分类效果以及计算复杂度,因此选取前 10 个奇异值进行归一化处理,作为图像的 SVD 特征。

## 4 遗传算法对特征集的优化组合

特征集优化组合是从原始特征中选择出一些有效性最高的特征以降低数据集维度的过程,由美国的 Holland<sup>[14]</sup> 提出。

借助玉米植株图像的颜色、SVD 及纹理特征提取了玉米植株的 20 维特征,其中,颜色特征 6 维、SVD 特征 10 维、纹理特征 4 维。为了获取最有效、最少的维数特征,本文利用 GA 对用于区分正常和特旱玉米植株图像的 20 维特征集进行了优化组合。基于 GA 的特征集组合优化算法流程如图 7 所示。

### 4.1 特征项个体表示和初始种群

GA 不能直接处理问题空间参数,必须对问题空间参数进行编码,由基因按照一定结构组成的染色体或者个体描述。本文采用二进制编码方法,即由二进制字符集  $\{0, 1\}$  将各个特征进行表示,基因 1 表示对应特征被选中,反之对应特征没被选中。例如若描述 6 个特征染色体的编码为 001010,则表示第 3 个和第 5 个特征被选中。随机产生 10 个染色体组成初始群体。

### 4.2 适应度确定和遗传操作

#### 4.2.1 适应度的定义

适应度用于评价算法过程中产生的个体的优劣程度。每次迭代过程中,根据适应度值对所有特征个体进行选择,以保证适应性好的个体有更大的几率繁殖后代,即优良基因(特征)得以遗传,因此适应度函数对于特征子集的寻优尤为关键。定义适应度函数为

$$F(X) = a - t \cdot \frac{p}{N}, \quad (6)$$

式中  $t$  为调整参数,  $X$  为所选中的特征子集,  $p$  为本次所选特征个数,  $N$  为原始特征个数,  $F$  为适应度函数,

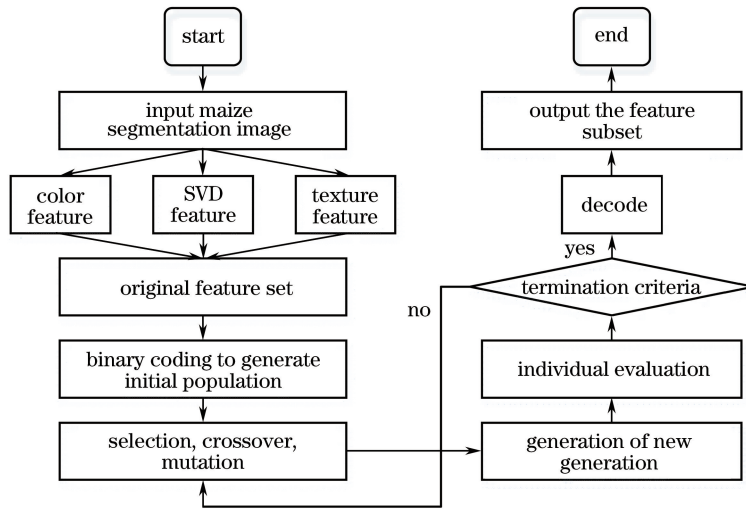


图7 特征组合优化算法流程图

Fig. 7 Flow chart of feature-set combinatorial optimization algorithm

$a$  为经过分类器所选中的特征组合的正确分类率。

$t$  用来平衡分类正确率和所选特征子集的比例,本文中选取  $t=0.5$ 。显然适应度值越大,表明所选的特征子集的性能越好,具体表现为特征个数少且分类正确率高。

#### 4.2.2 遗传算子的实现

GA中,染色体的进化在遗传算子中完成。遗传算子包括选择、交叉和变异。本文算法中选择算子为最优保存算子,概率为0.9,即保存上一代前90%的染色体;采用单点交叉法完成交叉操作,其概率为0.7;变异概率为0.02。

经过选择、交叉、变异完成了特征群体的遗传操作。每次迭代过程中经过遗传算子得到下一代群体,再计算每个个体的适应度。如果适应度满足要求,结束遗传操作,否则返回到遗传操作,直到运算到指定的最大迭代次数100。

通过该过程,最终降低了特征个数,完成了最优特征子集的选择。

## 5 玉米干旱识别实验研究结果

### 5.1 实验样本及分类器选择

正常植株图像152张,特旱植株图像180张,共计样本332张,图像大小均为640 pixel×480 pixel。随机取正常样本的2/3和特旱样本的2/3组成训练样本,剩余样本作为测试样本。

实验采用LSSVM作为分类器。该分类器在一定程度上降低了求解难度,提高了求解速度<sup>[15]</sup>。利用优化后的特征集合,重复玉米干旱识别实验100次,取正确识别率的均值作为最后的识别结果。

### 5.2 仿真结果对比

为了说明GA特征组合优化的性能,进行了如下玉米前期旱情的识别实验,实验结果如表2所示:

1) 将颜色、SVD、纹理分别作为单特征以及将20维特征直接融合作为分类器输入进行识别正确率的测试;

2) 采用GA将步骤1)中的20维特征进行优化,获取最优特征子集9维,包括2个颜色特征(G图和R图的偏度值)、5个SVD特征、2个纹理特征(相关性和逆差距),并将这9维特征作为分类器输入进行识别正确率的测试;

3) 选取对原有特征变量信息贡献率排名前9,且累计贡献率为95.23%的特征作为特征集合,采用主成分分析(PCA)法进行识别正确率测试。

表 2 不同特征在玉米图像旱情识别中平均识别正确率对比

Table 2 Comparison of average recognition accuracy of different features in maize image drought identification

Data name	Feature type	Feature number	Number of training samples	Number of test samples	Average accuracy	Average accuracy of normal plant	Average accuracy of seriously dry plant
Maize plant sample	Color	6	220 (normal plant 100, seriously drought sample 120)	112 (normal plant 52, seriously drought sample 60)	0.9503	0.9523	0.9485
	SVD	10			0.9627	0.9373	0.9847
	Texture	4			0.9771	0.9611	0.9852
	Color+SVD+texture	20			0.9460	0.9510	0.9417
	GA optimal feature	9			0.9903	0.9873	0.9928
	PCA	9			0.9745	0.9811	0.9702

由表 2 的结果可知,特征种类对干旱识别的正确率有很大的影响。颜色特征、SVD 特征和纹理特征均从不同方面反映了玉米图像的相关信息,单独使用某一类特征不能全面反映玉米干旱的情况,而直接将特征进行融合则可能会产生冗余,因此分类正确率相对较低;GA 相对于 PCA 法来说是一种特征的选择策略,能够在特征的不同组合之中选择比较好的解,选择的特征都是原始提取的特征,具有明确的意义。PCA 法是将特征进行组合,选择的特征是将原始提取的特征进行抽象处理得到的,同时依据识别结果,本文选择 GA 作为特征选择的策略。

### 5.3 分类器评价

正确率虽然是一个较为常用的指标,但并不能满足所有任务需求。对于分类器泛化性能进行评估,不仅需要有效可行的实验估计方法,还需要有衡量模型泛化能力的评价指标<sup>[16]</sup>。因此,为了进一步比较直接融合后的特征与经过 GA 选出的最优特征子集所训练出的分类器的性能,对两种不同的特征集合训练出的分类器的性能进行了测试。

混淆矩阵是清楚理解各种评价指标的基础,对于二分类问题可将样例参照其真实类别与学习器预测类别的组合划分成为真正例、假正例、真反例、假反例四种情形,分类结果的混淆矩阵如表 3 所示。根据混淆矩阵定义出查准率、查全率、真正率、假正率,其公式参照文献<sup>[17]</sup>。

表 3 分类结果混淆矩阵

Table 3 Classification results of confusion matrix

Actual label	Forecast label	
	Positive cases	Negative cases
Positive cases	Ture positive cases	False negative cases
Negative cases	False positive cases	Ture negative cases

以查准率作为纵轴,查全率作为横轴就得到了查准-查全(P-R)曲线,曲线越靠近右上角,说明分类器越能同时保证较高的查准率和查全率;以真正率作为纵轴,以假正率作为横轴,则得到受试者工作特征曲线(ROC),曲线越靠近左上角,说明分类器越能同时保证较高的真正率和假正率。通常情况下较为合理的判据是比较 ROC 曲线下的面积,即 AUC。图 8 表示两种特征训练出的分类器的 P-R 曲线以及 ROC 曲线,右

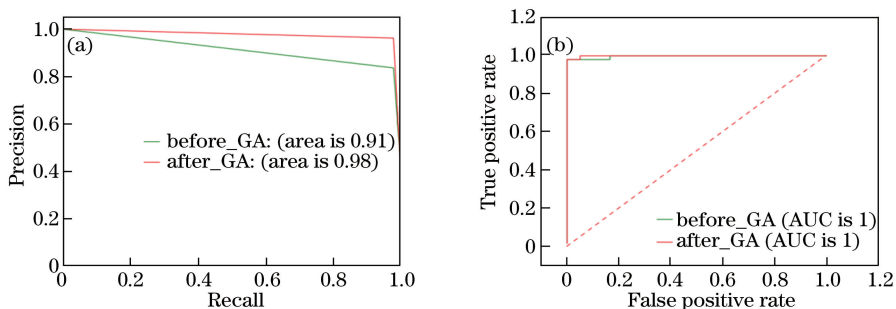


图 8 (a) P-R 曲线;(b) ROC 曲线

Fig. 8 (a) P-R curve; (b) ROC curve

下角给出曲线所包围的面积。从图 8 可以看出,直接融合后特征和优化后特征所训练出的分类器在 ROC 指标上表现基本相当,但在 P-R 曲线中,优化后的特征子集训练出的分类器的性能明显优于直接融合的特征所训练出的分类器的性能。说明经过 GA 得到的特征能够更加有效地构建 LSSVM,在各项指标上都有较好的表现。

## 6 结 论

与现有农业干旱监测方法相比,采用图像处理方法对玉米植株图像进行旱情识别具有一定的可行性,可以有效避免不利因素的干扰,有利于快速捕获旱情讯息,为农业旱情分析提供了新思路。

在算法方面,通过提取颜色、SVD 和纹理特征,获取了图像的不同特征,且特征之间具有一定的互补性,能够更加全面地对图像特征进行描述;采用 GA 对特征子集进行优化。

从识别结果可看出,该方法对正常和特旱植株具有较好的分类效果,且对于不同位置、有无遮挡物的图像均具有较强的识别能力。

农业上,干旱情况分类特旱、重旱、中旱、轻旱和正常,如何提取更加有效的图像特征,训练对不同于旱情况下都有较强识别能力的分类器是下一步工作的重点。

## 参 考 文 献

- [1] Li Maosong, Li Sen, Li Yuhui. Studies on drought in the past 50 years in China [J]. Chinese Journal of Agrometeorology, 2003, 24(1): 7-10.  
李茂松, 李 森, 李育慧. 中国近 50 年旱灾灾情分析[J]. 中国农业气象, 2003, 24(1): 7-10.
- [2] Ni Shenhai, Gu Ying, Yan Nana, *et al.* Research on agricultural drought assessment method based on information assimilation of moisture cycle simulation model and remote sensing image information [J]. China Rural Water and Hydropower, 2016(1): 51-54.  
倪深海, 顾 颖, 闫娜娜, 等. 农业旱情评估方法研究——基于农田水分循环模拟与遥感影像信息同化[J]. 中国农村水利水电, 2016(1): 51-54.
- [3] Lu H, Cao Z, Xiao Y, *et al.* Fine-grained maize tassel trait characterization with multi-view representations [J]. Computers and Electronics in Agriculture, 2015, 118: 143-158.
- [4] Camargo A, Smith J S. Image pattern classification for the identification of disease causing agents in plants [J]. Computers and Electronics in Agriculture, 2009, 66(2): 121-125.
- [5] Selim S Z, Ismail M A. K-means-type algorithms: A generalized convergence theorem and characterization of local optimality[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1984, 6(1): 81-87.
- [6] Lu Zhiying, Liu Hai, Jia Huizhen, *et al.* Recognition of hail and rainstorm based on the radar reflectivity image features[J]. Acta Physica Sinica, 2014, 63(18): 189201.  
路志英, 刘 海, 贾惠珍, 等. 基于雷达反射率图像特征的冰雹暴雨识别[J]. 物理学报, 2014, 63(18): 189201.
- [7] Pan Haibin, Zhang Wei, Cong Mingyu, *et al.* Image preprocessing algorithm for space target space-based detection[J]. Acta Optica Sinica, 2009, 29(9): 2402-2407.  
潘海斌, 张 伟, 丛明煜, 等. 一种空间目标在轨检测图像预处理算法[J]. 光学学报, 2009, 29(9): 2402-2407.
- [8] Wu J Y. A image retrieval method based on color histogram[C]. Informatics in Control, Automation and Robotics, 2012, 133: 741-746.
- [9] Haralick R M, Shanmugam K, Dinstein I. Texture features for image classification [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1973, 3(6): 610-621.
- [10] Liu Tao, Sun Chengming, Wang Lijian, *et al.* In-field wheatear counting based on image processing technology [J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(2): 282-290.  
刘 涛, 孙成明, 王力坚, 等. 基于图像处理技术的大田麦穗计数[J]. 农业机械学报, 2014, 45(2): 282-290.
- [11] Zhang S W, Wang Z. Cucumber disease recognition based on global-local singular value decomposition [J]. Neurocomputing, 2016, 205: 341-348.
- [12] Chang Liping, Shen Weixing, Lin Zunqi. Algorithm for digital wavefront fitting based on singular value decomposition



- [J]. *Acta Optica Sinica*, 2006, 26(11): 1676-1680.
- 常丽萍, 沈卫星, 林尊琪. 基于奇异值分解的数字波前拟合法[J]. *光学学报*, 2006, 26(11): 1676-1680.
- [13] Zhang Xianda. *Matrix analysis and applications*[M]. Beijing: Tsinghua University Press, 2013.
- 张贤达. 矩阵分析与应用[M]. 北京: 清华大学出版社, 2013.
- [14] Holland J H. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, artificial intelligence*[M]. Cambridge: MIT Press, 1992.
- [15] Chen Xiaofeng, Long Changjiang, Niu Zhiyou, *et al.* Classification research of Chinese medicine based on latent semantic analysis and NIR[J]. *Acta Optica Sinica*, 2014, 34(9): 0930001.
- 陈晓峰, 龙长江, 牛智有, 等. 基于潜在语义分析与 NIR 的中药材分类研究[J]. *光学学报*, 2014, 34(9): 0930001.
- [16] Fawcett T. An introduction to ROC analysis[J]. *Pattern Recognition Letters*, 2006, 27(8): 861-874.
- [17] Zhou Zhihua. *Machine learning*[M]. Beijing: Tsinghua University Press, 2016: 28-37.
- 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 28-37.