

## 模糊数学在烟叶分级中的应用\*

张 帆 张新红 张 彤

## 摘 要

根据国家颁布的烤烟 40级国家标准,烤烟分为 40个级别。由于国家标准仅对烟叶的外观质量进行了定性描述,因此评判烟叶的质量存在很大的主观性和模糊性。模糊数学是一门研究和处理现实世界中广泛存在的一类模糊现象的学科。在应用计算机图像处理技术的基础上,利用模糊数学和模式识别的技术,将有可能实现烟叶的自动分级。试验表明,应用模糊数学的方法,可以较好的解决烟叶自动分级问题。

关键词: 模糊数学 烟叶

中图分类号: TS46 文献标识码: B 文章编号: 1004-5708(2002)03-0044-05

目前我国烟叶的品质检测分级技术还十分落后,绝大部分还停留在依靠人工感官判断的原始阶段。这种主观判定受到人的视力、颜色鉴别能力、情绪、疲劳和光线条件等的限制,效率低,误差大。根据国家颁布的烤烟 40级国家标准,烤烟分为 40个级别。由于国家标准仅对烟叶的外观质量进行了定性描述,因此评判烟叶的质量存在很大的主观性和模糊性。用人工感官对烟叶的品质进行评判,一般是从颜色、大小、性状、表面状态、表面缺陷等几个方面进行评判,绝大部分特征与人的视觉有关。随着计算机技术的迅猛发展,近二、三十年来才形成和发展起来的计算机图像处理技术,机器视觉,模式识别以及模糊数学在模拟人的视觉方

面取得了重大的进展。在应用计算机图像处理技术的基础上,利用模糊数学和模式识别实现烟叶的自动分级,将有助于改进我国烟叶分检手段落后的局面。

## 1 提取烟叶的特征值

## 1.1 图像采集系统

图像采集系统的硬件组成如图 1所示。

将待测烟叶放置于光源箱内,打开 CCD摄像头利用我们编写的软件即可拍摄图像。所拍摄的烟叶图像可显示、保存在计算机上。拍摄图像的尺寸是 512×512,真彩色图像。

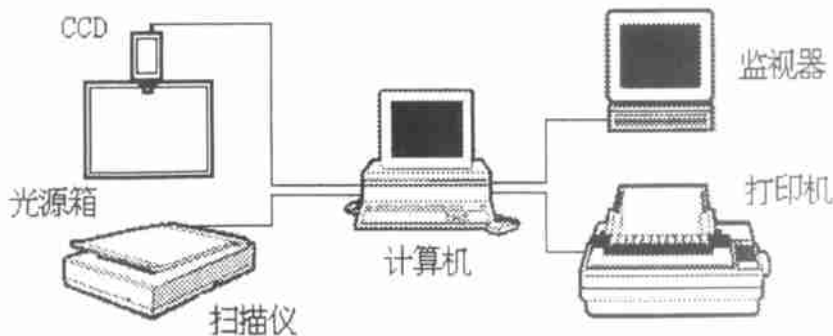


图 1 图像采集系统示意图

\* 张帆,男,34岁,博士研究生,北京工业大学计算机学院计算机科学学科部,北京朝阳区平乐园 100号,北京,100022

张新红,张彤,通讯地址同第一作者

收稿日期: 2001-03-13

CCD摄像头: Panasonic WV-CL500 CCD真彩色  
图像传感器

图像采集卡: CPE-1000真彩色图像采集卡。  
光源箱: 尺寸 0.66× 0.66× 0.56m, 内置 4支 18W日光灯。

1.2 提取烟叶的特征值

拍摄烟叶图像后,我们利用图像处理技术提取了烟叶图像的形状特征、颜色特征和纹理特征共 29个,存储在数据库中。利用我们编写的软件,处理一幅 512 × 512真彩色烟叶图像,得到这些烟叶特征值约需 10s左右的时间

提取的烟叶特征值分别是:  
形状特征: 烟叶的面积、周长、长度、宽度、缺损、长宽比、圆弧度。

颜色特征: 红色通道均值、绿色通道均值、蓝色通道均值、红色通道方差、绿色通道方差、蓝色通道方差、红色通道峰度、绿色通道峰度、蓝色通道峰度、色度均值、饱和度均值、亮度均值、色度方差、饱和度方差、亮度方差、色度峰度、饱和度峰度、亮度峰度

纹理特征: 图像的纹理熵、能量、惯性、逆差距  
1.2.1 形状特征值提取 要提取烟叶的形状特征值首先需要得到烟叶的边缘数据。图像边缘检测可以采用拉普拉斯 (Laplacian)算子,它定义为:

$$\nabla^2 f(x,y) = \frac{\partial^2}{\partial x^2} f(x,y) + \frac{\partial^2}{\partial y^2} f(x,y)$$

以差分运算来代替微分运算,则:

$$\nabla^2 f(x,y) = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y)$$

由于拉普拉斯算子是一个二阶导数,二阶导数在图像的边缘处为零,它将在边缘处产生一个陡峭的零交叉。但是,试验表明由于噪声的原因,采用拉普拉斯算子进行边缘检测效果不理想,需要采用高斯 (Gaussian)低通滤波消除噪声。

高斯滤波表达式为:

$$g(x,y) = G(x,y) * f(x,y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) * f(x,y)$$

式中  $\sigma$  为高斯函数的标准差  
\* 为卷积  
可以把拉普拉斯算子和高斯脉冲响应组合为一个高斯拉普拉斯算子:

$$\nabla^2 \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) = -\frac{1}{\sigma^4} \left(1 - \frac{x^2+y^2}{2\sigma^2}\right) \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

$\left(-\frac{x^2+y^2}{2\sigma^2}\right)$   
对经过高斯拉普拉斯算子滤波后的图像用零灰度值进行二值化会产生闭合、连通的轮廓并消除了所有的内部点

得到烟叶的边缘数据后,对这个边界曲线进行多形拟合。计算拟合多边形各边边长之和作为烟叶的周长,整个拟合多边形的面积作为烟叶的面积。拟合多边形的面积可由下式得到:

$$A = \frac{1}{2} \sum_{i=1}^{N_b} (x_i y_{i+1} - x_{i+1} y_i)$$

式中:  $N_b$  拟合多边形界点的数目

同样,可以用拟合多边形的长度和宽度来近似烟叶的长度和宽度

烟叶缺损的计算方法是,扫描整幅图像,记录所有在烟叶边界之内但颜色为背景色的像素点数,以及所有在烟叶边界之内的像素点数,取二者之比的百分数为烟叶的缺损。另外,利用这个拟合多边形也可以计算出烟叶的叶尖角

1.2.2 颜色特征值提取 我们可以利用图像处理的方法得到烟叶图像中各像素点处的颜色值,由于提取或计算的图像颜色特征是离散数据,因此很容易计算它们的概率分布。经常使用且具有很好表述能力的特征有: 均值、方差、偏度、峰度、能量等,这些特征在离散情况下的计算公式详见作者的另外两篇论文<sup>[1,2]</sup>。

1.2.3 纹理特征值提取 空间灰度共生矩阵 (Grayscale Co-occurrence Matrix)方法是建立在估计图像的二阶组合条件概率密度函数基础上的。在纹理图像中,在某个方向上相隔一定距离的一对像素灰度出现的统计规律,能具体反映这个图像的纹理特征。可以用灰度共生矩阵来描述这个统计规律,进而由灰度共生矩阵计算出一些参数定量描述这些纹理特征

灰度共生矩阵被定义为从灰度值为  $i$  的像素点离开某个固定位置关系  $\delta = (DX, DY)$  的点上的灰度为  $j$  的概率或频率,即:

设图像的大小为  $W \times L$  像素,则:

$$p^w(i,j) = \{(x,y) | f(x,y) = i, \text{ and } f(x+DX, y+DY) = j\}$$

$x, y = 0, 1, 2, \dots, W-1$   
 $i, j = 0, 1, 2, \dots, N-1$   
式中:  $x, y$  图像中像素坐标;

$f(x,y)$ 像素灰度;

$N$  灰度级的数目。

一幅图像的灰度共生矩阵可以表示为:

$$p_{\delta}(i,j)=\begin{bmatrix} p_{\delta}(0,0) & \delta & p_{\delta}(0,N-1) \\ \delta & p_{\delta}(i,j) & \delta \\ p_{\delta}(N-1,0) & \delta & p_{\delta}(N-1,N-1) \end{bmatrix}$$

我们提取的烟叶纹理特征值提取是利用烟叶图像的灰度共生矩阵求其数字统计特征,计算公式分别如下:

纹理熵:

$$H=\sum_{i=1}^N\sum_{j=1}^NP_{ij}\log P_{ij}$$

纹理能量:

$$E=\sum_{i=1}^N\sum_{j=1}^N(P_{ij})^2$$

纹理惯性:

$$F=\sum_{i=1}^N\sum_{j=1}^N(i-j)^2P_{ij}$$

纹理逆差矩:

$$L=\sum_{i=1}^N\sum_{j=1}^N\frac{1}{1+(i-j)^2}P_{ij}$$

式中:  $P$  经归一化处理后的灰度共生矩阵元素  
 $N$  图像的灰度级数目

## 2 利用模糊数学技术实现烟叶的自动分级

我们提取的烟叶特征值,即模糊数学中的因素多达 29个。在下面的试验中,我们只提取了样本烟叶的 8个特征值,而且假定样本烟叶只可能属于 3个级别而不是 40个级别。这 3个级别分别是下部柠檬黄 1级,上部柠檬黄 4级和光滑叶 1级。

需要说明的是,烟叶的自动分级是一个十分复杂的问题,国标关于分级指标的规定多为定性的,如成熟度、身份、油份等很难以定量的数据来描述,对此我们进行了大量的研究,提取了 29个烟叶特征值。这里我们选取 8个特征值进行试验只是为了说明利用模糊数学技术实现烟叶的自动分级的可行性,如果进行进一步的试验,不但要利用上述 29个特征值,还要做深入的研究,以提取出更多、更有效的特征值。

我们选取的待评判烟叶样本特征值即因素如表 1

| 表 1 烟叶样本特征值 |     |       |       |      |       |      |     |     |
|-------------|-----|-------|-------|------|-------|------|-----|-----|
|             | 周长  | 面积    | 缺损    | 纹理能量 | 纹理熵   | 红方差  | 绿方差 | 蓝方差 |
| 样本烟叶        | 820 | 18326 | 0.038 | 0.72 | -0.89 | 1032 | 853 | 698 |

表中: 周长、面积的单位为像素数,缺损的单位为%,其余特征值无量纲。

2.1 模糊综合评判

烟叶样本中,同一级别中同一特征值的均值分别如表 2

3个级别的标准样本特征值,即在一批标准

| 表 2 三个级别的烟叶标准样本特征值 |     |       |       |      |       |      |     |     |
|--------------------|-----|-------|-------|------|-------|------|-----|-----|
|                    | 周长  | 面积    | 缺损    | 纹理能量 | 纹理熵   | 红方差  | 绿方差 | 蓝方差 |
| 下部柠檬黄              | 212 | 41937 | 0.011 | 0.48 | -1.43 | 887  | 696 | 575 |
| 1级                 | 3   |       |       |      |       |      |     |     |
| 上部柠檬黄              | 941 | 16478 | 0.036 | 0.71 | -0.90 | 1027 | 826 | 593 |
| 4级                 |     |       |       |      |       |      |     |     |
| 光滑叶                | 206 | 41242 | 0.080 | 0.42 | -1.53 | 633  | 826 | 593 |
| 1级                 | 9   |       |       |      |       |      |     |     |

表中: 周长、面积的单位为像素数,缺损的单位为%,其余特征值无量纲。

由于影响烟叶分级的因素很多,我们将采取二级模糊综合评判的方法 具体的评判步骤如下:

(1) 将因素分类

我们把影响烟叶分级的因素分为 3类,即将因素集  $U$  分为 3个因素子集:

$$U = \{\text{外观特征, 纹理特征, 颜色特征}\}$$

再考虑每个因素子集中的因素如下:

$$U_1 = \{\text{周长, 面积, 缺损}\}$$

$$U_2 = \{\text{纹理能量, 纹理熵}\}$$

$$U_3 = \{\text{红方差, 绿方差, 蓝方差}\}$$

(2) 建立权重集

① 因素类权重集

根据我们的试验,在颜色、形状、纹理 3类特征值中,最有效的是颜色特征值,可以有效的识别烟叶的含青面积与青色程度的不同 利用下部烟叶叶形圆宽,上部烟窄长,中部烟适中的特征,周长、面积、缺损特征值也会对识别烟叶有一定帮助 烟叶的纹理特征对烟叶的识别作用较小

所以,根据各类因素的重要程度的不同,赋予每个因素类以相应的权重 上述 3个因素类的权重如下:

$$A = (0.35, 0.2, 0.45)$$

② 因素权重集

在各类因素中,经过大量试验,根据各因素的重要程度的不同,赋予每个因素以相应的权重 因素权重集如下:

$$A_1 = (0.3, 0.4, 0.3)$$

$$A_2 = (0.5, 0.5)$$

$$A_3 = (0.3, 0.3, 0.4)$$

③ 建立评判集

我们假定待评判的烟叶只可能属于 3个级别,则评判集可表示为:

$$V = \{\text{下部柠檬黄 1级, 上部柠檬黄 4级, 光滑叶 1级}\}$$

④ 一级模糊综合评判

按一类中的各个因素进行综合评判,则一级模糊综合评判的单因素评判矩阵为:

$$R_1 = \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.6 & 0.1 \end{bmatrix}$$
$$R_2 = \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

$$R_3 = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.7 & 0.2 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$

于是,3类因素的模糊综合评判集分别如下:

$$B_1 = A_1 \circ R_1 = (0.3, 0.4, 0.3) \circ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.6 & 0.1 \end{bmatrix}$$
$$= (0.3, 0.4, 0.3)$$

$$B_2 = A_2 \circ R_2 = (0.5, 0.5) \circ \begin{bmatrix} 0.2 & 0.5 & 0.3 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$
$$= (0.3, 0.5, 0.3)$$

$$B_3 = A_3 \circ R_3 = (0.3, 0.3, 0.4) \circ \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.1 & 0.7 & 0.2 \\ 0.3 & 0.5 & 0.2 \end{bmatrix}$$
$$= (0.3, 0.4, 0.2)$$

⑤ 二级模糊综合评判

二级模糊综合评判中的单因素评判矩阵为一级模糊综合评判矩阵:

$$R = \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.3 & 0.5 & 0.3 \\ 0.3 & 0.4 & 0.2 \end{bmatrix}$$

于是,二级模糊综合评判集如下:

$$B = A \circ R = (0.35, 0.2, 0.45) \circ \begin{bmatrix} 0.3 & 0.4 & 0.3 \\ 0.3 & 0.5 & 0.3 \\ 0.3 & 0.4 & 0.2 \end{bmatrix}$$
$$= (0.3, 0.4, 0.3)$$

上述的二级模糊综合评判集表明,这个待评判烟叶对下部柠檬黄 1级的隶属度为 0.3,对上部柠檬黄 4级的隶属度为 0.4,对光滑叶 1级的隶属度为 0.3

2.2 模式识别检验

下面我们对上述的二级模糊综合评判结果用模式识别的方法进行检验

统计表明,烟叶各特征值即模糊数学中的因素呈正态分布,即隶属函数为正态分布 首先利用统计方法确定 3个级别中各特征值的均值和方差 其中 3个级别特征值的均值如表 2所示 然后根据各特征值的均值和方差,求待评判烟叶样本各特征值在 3个级别中对应特征值的正态分布值,作为该特征值对该级别的隶属度 如表 3所示

由表可得:

$$u(\text{下部柠檬黄 1级}) = 0.3,$$

$$u(\text{上部柠檬黄 4级}) = 0.7,$$

$$u(\text{光滑叶 1级})= 0.3$$

$$\max(0.3, 0.7, 0.3)= 0.7= u(\text{上部柠檬黄 4级})$$

表 3 3个级别烟叶样本特征值的隶属度

|          | 周长  | 面积  | 缺损  | 纹理能量 | 纹理熵 | 红方差 | 绿方差 | 蓝方差 |
|----------|-----|-----|-----|------|-----|-----|-----|-----|
| 下部柠檬黄 1级 | 0.2 | 0.2 | 0.3 | 0.2  | 0.3 | 0.3 | 0.1 | 0.3 |
| 上部柠檬黄 4级 | 0.5 | 0.6 | 0.6 | 0.5  | 0.5 | 0.5 | 0.7 | 0.5 |
| 光滑叶 1级   | 0.3 | 0.2 | 0.1 | 0.3  | 0.2 | 0.2 | 0.2 | 0.2 |

表中: 周长、面积的单位为像素数,缺损的单位为%,其余特征值无量纲。

所以,根据最大隶属度原则,把这个烟叶样本归属于上部柠檬黄 4级。这也印证了二级模糊综合评判的结果

3 结论

从上述试验可以看出,应用模糊数学的方法,有可能较好地解决烟叶的自动分级问题。但是由于烟叶本身的复杂性,还需要进行大量的试验,摸索出规律性,争取能利用计算机图像处理技术和模糊数学相结合的新算法改进我国烟叶分检手段落后的局面。

参考文献

1 张帆,方如明,蔡健荣. 基于神经网络提取烟叶重量特征值的研究. 农业机械学报,2000(6).

2 张帆,方如明,蔡健荣. 标准烟叶数据库的图像检索. 农业机械学报,2001(1).

3 王润生. 图像理解. 长沙: 国防科技大学出版社,1995.

4 Castleman K R(朱志刚等译). 数字图像处理. 北京: 电子工业出版社,1998.

5 黄凤岗,宋克欧. 模式识别. 哈尔滨: 哈尔滨工程大学出版社,1998

6 戚飞虎,等. 模式识别与图像处理. 上海: 上海交通大学出版社,1990.

7 张世伟,等. 模糊数学应用. 上海: 同济大学出版社,1991.

8 模糊技术与神经网络技术选编. 北京: 北京航空航天大学出版社,1999.

9 王士同. 模糊推理与模糊专家系统. 上海: 上海科学技术文献出版社. 1994.

10 中国烟草工业总公司郑州烟草研究院. 烤烟国家标准培训教材,1992.

Application of fuzzy mathematics in the classification of tobacco leaves

Zhang Fan Zhang Xinhong Zhang Tong  
Beijing Polytechnic University, Computer Institute, Beijing 100022

Abstract

According to the new flue-cured tobacco state standard leaves can be graded into 40 classes. Because state standard only describes the tobacco appearance, so judging tobacco quality will be subjective and fuzzy. Fuzzy mathematics is a subject about studying and dealing with many fuzzy phenomena in our practical world. On the base of application of computer image processing, using fuzzy mathematics and pattern recognition technology, tobacco grading can be automated. The experiment shows that fuzzy mathematics method can preferably resolve auto-tobacco grading problem.

**Key words** Fuzzy mathematics Tobacco leaf Grading